PL-TR-96-2074

ADVANCED GEOPHYSICAL ENVIRONMENT SIMULATION TECHNIQUES

D.B. Hogan
G.B. Gustafson
R.P. d'Entremont
C.F. Ivaldi
C.P. Sarkisian

Atmospheric and Environmental Research, Inc.
840 Memorial Drive
Cambridge, MA 02139

8 March 1996

Scientific Report No. 1

Approved for public release; distribution unlimited

PHILLIPS LABORATORY
Directorate of Geophysics
AIR FORCE MATERIEL COMMAND
HANSCOM AFB, MA 01731-3010

"This technical report has been reviewed and is approved for publication"

EDWARD C. ROBINSON
Contract Manager
Data Analysis Division

ROBERT E. McINERNEY, Director
Data Analysis Division

This report has been reviewed by the ESD Public Affairs Office (PA) and is releasable to the National Technical Information Service (NTIS).

Qualified requestors may obtain additional copies from the Defense Technical Information Center. All others should apply to the National Technical Information Service.

If your address has changed, or if you wish to be removed from the mailing list, or if the addressee is no longer employed by your organization, please notify PL/IM, 29 Randolph Road, Hanscom AFB, MA 01731-3010. This will assist us in maintaining a current mailing list.

Do not return copies of this report unless contractual obligations or notices on a specific document requires that it be returned.

REPORT DOCUMENTATION PAGE	*Form Approved OMB No. 0704-0188*

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	8 March 1996	Scientific Report No. 1

4. TITLE AND SUBTITLE

Advanced Geophysical Environment Simulation Techniques

5. FUNDING NUMBERS

CONTRACT F19628-94-C-0106
PE 35160F
PR 7659 TA GY WU AC

6. AUTHOR(S)

D.B. Hogan, G.B. Gustafson, R.P. d'Entremont, C.F. Ivaldi, C.P. Sarkisian

7. PERFORMING ORGANIZATION NAMES(S) AND ADDRESS(ES)

Atmospheric and Environmental Research, Inc.
840 Memorial Drive
Cambridge, MA 02139

8. PERFORMING ORGANIZATION REPORT NUMBERS

9. SPONSORING / MONITORING AGENCY NAMES(S) AND ADDRESS(ES)

Phillips Laboratory
29 Randolph Road
Hanscom AFB, MA 01731-3010
Contract Manager: Edward C. Robinson/GPD

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

PL-TR-96-2074

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for public release; distribution unlimited

12b. DISTRIBUTION CODE

13 ABSTRACT *(Maximum 200 words)*

This report describes a multi-faceted effort to support PL/GPD in the development and application of state-of-the-art analysis techniques for remotely sensed data. The resulting tools, techniques, and data sets support the improved analysis of archived data as well as current and future geophysical parameter acquisition and analysis.

The following tasks are addressed: (1) development of cloud data sets to support cloud forecast model development; (2) further enhancements to cloud analysis algorithms and development of data sets to support several DoD infrared sensing missions; and (3) upgrades to the Air Force Meteorological System (AIMS) to add GOES-8 reception capability and an improved relational database.

14. SUBJECT TERMS	15. NUMBER OF PAGES
clouds, satellite meteorology, remote sensing, database	234
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Unlimited

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION AND OVERVIEW

The Air Force Phillips Laboratory Data Analysis Division (PL/GPD), owing to its long history of earth, atmospheric and ionospheric measurement programs, has collected an abundance of data in a variety of physical regimes. A lasting result of this activity is that a variety of geophysical data sets, obtained at various times during the past and with vastly different instrumental and computing technologies, could now be reanalyzed and applied to a variety of current programs. One of the major difficulties encountered in accessing a specialized database centers on the need to decode and translate the data set into an acceptable format for further processing.

This document describes a multi-faceted effort to support PL/GPD in the development and application of state-of-the-art analysis techniques for remotely sensed data. The resulting tools, techniques, and data sets will support the improved analysis of archived data as well as current and future geophysical parameter acquisition and analysis. The principal goals of this effort are to:

- develop and enhance data access methods for acquiring and processing geophysical data sets;

- refine and optimize advanced satellite data analysis software for application to the above data sets; and

- prepare and analyze coordinated and quality controlled data sets of raw and processed satellite data for use in simulation and research.

Specifically, this project is centered around the Air Force Interactive Meteorological System (AIMS) located at the Air Force Phillips Laboratory Atmospheric Sciences Division (PL/GPA) facility at Hanscom Air Force Base. AIMS is an integrated facility consisting of multiple real-time meteorological satellite receiving stations, a comprehensive archive of satellite and conventional data products, a powerful database access system, a rich set of state-of-the-art algorithms for satellite data analysis, and a mixed computing environment. It is operated jointly by the US Air Force and Atmospheric and Environmental Research, Inc. (AER) under the terms of a Cooperative Research and Development Agreement (CRADA).

One of the key AIMS assets is its catalog of state-of-the-art satellite cloud analysis algorithms. These algorithms have been developed by AER during the past eight years under Phillips Laboratory sponsorship. The key research project codes are:

- TACNEPH (contract # F19628-90-C-0112): basic cloud properties (cover, layers, type) for DMSP and NOAA satellites as appropriate for processing in a military tactical environment;

- SERCAA Phase I (contract # F19628-92-C-0149): basic cloud properties (cover, layers, type) for both polar-orbiting (DMSP and NOAA series) and geosynchronous (GOES, GMS and METEOSAT) satellites plus a merge processing step which produces a best synoptic analysis from all available data;

- SERCAA Phase II (contract # F19628-92-C-0149): enhanced cloud properties including cirrus properties (emissivity, effective height and effective temperature), cloud phase, and integration of cloud analysis with microwave sounding.

The SERCAA Phase I algorithms (Gustafson et al., 1994) are currently in the process of being implemented operationally at Air Force Global Weather Central (AFGWC) as part of the Cloud Depiction and Forecast System II program (CDFS II). The algorithms have also been selected by the CERES and MODIS Science Teams as the first-step cloud clearing algorithms for use in the processing of data from these instruments (Moderate Resolution Imaging Spectroradiometer (MODIS) and Clouds and Earth's Radiant Energy System (CERES)). The SERCAA Phase II (d'Entremont et al., 1996) algorithms for cirrus property determination provide extremely important cloud parameters that are not available through other available cloud algorithms (e.g., Clouds from AVHRR (CLAVR) and International Satellite Cloud Climatology Project (ISCCP)). AIMS also hosts a large quantity of satellite data, obtained both from the AIMS satellite ground stations and from external sources. The combination of these capabilities makes AIMS an extremely valuable resource for generating analyzed cloud data to support a variety of programs.

Under this project, we have used AIMS to generate raw and analyzed cloud data sets to support two major programs. The first task is the generation of data sets to support the development of advanced worldwide cloud forecast models by the Defense Nuclear Agency (DNA). This project employed only the SERCAA Phase I algorithms. The second task was to generate data sets to support the Space Based Infrared System (SBIRS) backgrounds Phenemonology Science Working group. This application required both the SERCAA Phase I algorithm results and the cirrus retrievals of SERCAA Phase II.

In conjunction with the above tasks, AER has continued to support ongoing maintenance of AIMS. In addition, AER is undertaking several tasks to upgrade and improve AIMS:

- Upgrade of geosynchronous satellite reception capability to include GOES-8 (and preparations/analysis for a GOES-9 capability)
- Initiation of a major upgrade to the AIMS database, including integration of a new server hosting a relational database management system from Oracle
- Recoding/rehosting selected algorithms to facilitate generation of the large scale satellite data sets described above

This report is organized as follows. Section 2 summarizes the work on the DNA Worldwide Satellite Data Set generation. Section 3 describes the upgrade of the satellite receiving station to include a GOES-8 capability. Section 4 presents the work on the data set generation to support the SBIRS Phenemonology Science Working Group. Sections 5 concludes with a summary of the status of the AIMS database upgrade and related efforts. Appendices contain supporting documentation prepared with the support of this contract. Appendix A contains Data Save Document Reports that were previously submitted in support of the DNA Worldwide Satellite Data Sets project. These reports contain detailed information describing the data sets themselves as well as their format, how they were gathered and processed, and descriptions of the algorithms used to generate them. Appendix B contains a copy of a presentation concerning the AIMS Database Upgrade.

2. DNA WORLDWIDE SATELLITE DATA SETS

Substantial changes in the operational satellite cloud analysis methodology in use at the Air Force Global Weather Central (AFGWC) will occur as part of the Cloud Depiction and Forecast System II (CDFS II) development effort. The existing cloud analysis model, known as RTNEPH (Real-Time NEPHanalysis), will be replaced by a new model based on a series of algorithms developed by AER under contract to the Air Force Phillips Laboratory (PL) through support from the Strategic Environmental Research and Development Program (SERDP). The new cloud analysis model is referred to as the Support of Environmental Requirements for Cloud Analysis and Archive (SERCAA) model (Gustafson et al., 1994). In addition to cloud analysis, CDFS II will also implement a new cloud forecast model. However, the CDFS II forecast model is a combination of previously existing models and contains no new science. The Defense Nuclear Agency (DNA) is conducting an independent research effort to develop and test a new cloud forecast model capable of generating both short and long-term forecasts that is potentially applicable to CDFS II. To support the forecast model development effort, DNA funded work under this contract to generate SERCAA data products for use in testing the new forecast models. SERCAA data products were selected to provide compatibility with the cloud analysis products that will eventually be produced by CDFS II.

SERCAA cloud analysis algorithms provide for integration of sensor data from both military and civilian polar orbiting environmental satellites plus high temporal resolution imagery from geostationary platforms. AER was responsible for all development and testing of SERCAA cloud analysis algorithms under a basic research contract with PL and, as such, was able to provide unique insight into the characteristics of the CDFS II cloud model. A major innovation of the SERCAA cloud analysis algorithms is the use of high spatial, spectral, and temporal resolution multispectral data obtained from multiple satellite systems. By contrast, RTNEPH is constrained to operate only on polar orbiting satellite data maintained at degraded spatial and spectral resolution in the AFGWC Satellite Global DataBase (SGDB). SGDB is limited to single visible and infrared channel data from either DMSP or, when DMSP data are unavailable, NOAA polar orbiting operational satellites. Further, AFGWC employs a remapping process to warp the sensor data to a standardized polar stereographic map projection with a concomitant reduction in spatial resolution from 2.7 km to approximately 6 km. Radiometric resolution is also reduced from 8 (DMSP) or 10 (NOAA) bits to 6 bits (i.e., 64 discrete values), resulting in a thermal resolution of infrared brightness temperatures of approximately 1.9 K.

SERCAA algorithms operate on calibrated sensor data at the full spatial and spectral resolution of the sensor (Table 1). In addition to two channel DMSP/OLS data, multispectral NOAA/AVHRR and hourly geostationary data from GOES, GMS, and METEOSAT, are also processed. Data from the recently launched US GOES I-M series and the proposed Chinese Feng Yun 2A satellites can also be accommodated with minor modifications to the algorithms. To best exploit the information content from each sensor, and to minimize distortion of the data, cloud analysis is performed in the raw satellite scan projection using all available data bits. No remapping or truncating of the pixel data occurs. To accommodate this approach, sensor data from the individual satellite systems are analyzed separately using multiple analysis algorithms, each designed to exploit the unique sensor data attributes of a particular satellite system. The analysis algorithms are organized into four processing layers as illustrated in Figure 1. The first layer employs data ingest code for each satellite system. The second consists of

Table 1. Sensor Channel Data Attributes Used for DNA Study

Satellite	Sensor	Channel (µm)	Data Format	Resolution[1] (km)	Bits per Pixel[2]	Pixels per Scan Line
DMSP	OLS	0.40-1.10	counts	2.7	6	1464
		10.5-12.6	EBBT	2.7	8	1464
NOAA	AVHRR	0.58-0.68	percent albedo	4.0	8	409
		0.72-1.10	percent albedo	4.0	8	409
		3.55-3.93	EBBT	4.0	8	409
		10.3-11.3	EBBT	4.0	8	409
		11.5-12.5	EBBT	4.0	8	409
GOES	VAS	0.55-0.75	counts	0.9	6	15288
		3.71-4.18	EBBT	14.0	8	1911
		10.5-12.6	EBBT	7.0	8	3822[3]
METEOSAT	VISSR	0.55-0.75	counts	2.5	8	5000
		10.5-12.6	EBBT	5.0	8	2500
GMS	VISSR	0.5-0.75	counts	1.25	6	10000
		10.5-12.5	EBBT	5.0	8	2500

[1] sensor resolution at satellite subpoint

[2] AVHRR radiance data are transmitted at 10-bit resolution, however, the SERCAA development system could only accommodate 8-bit brightness temperature data (although the full 10-bit resolution is used in the radiance to brightness temperature transformation)

[3] GOES long wave infrared data are over sampled in the across-track direction by a factor of 2.

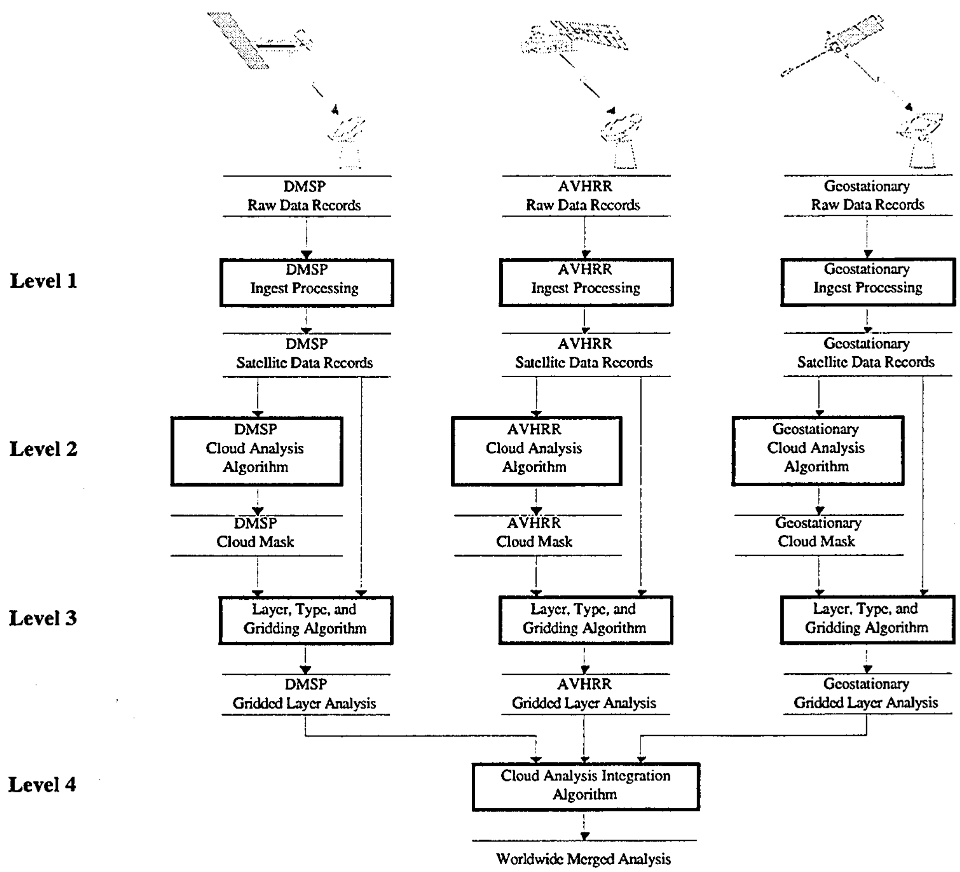


Figure 1. SERCAA/CDFS II-Nephanalysis Processing Flow

the three satellite specific analysis algorithms, one each for OLS, AVHRR, and geostationary data. These algorithms provide cloud location information on a pixel-by-pixel basis. The third layer further analyzes regions classified as cloud to provide information on the vertical distribution of cloud layers including number, height, and type. Cloud layer information is accumulated over a standard AFGWC 16th mesh polar grid and layer cloud amounts along with total cloud amount are computed for each grid box. The final processing layer analyzes total cloud and layer information derived from the separate satellite specific algorithms to produce a single integrated cloud analysis. The integration algorithm uses a rules based approach combined with a modified optimal interpolation scheme to account for differences in timeliness and accuracy characteristics in the separate, asynchronous cloud analyses produced by the earlier processing layers.

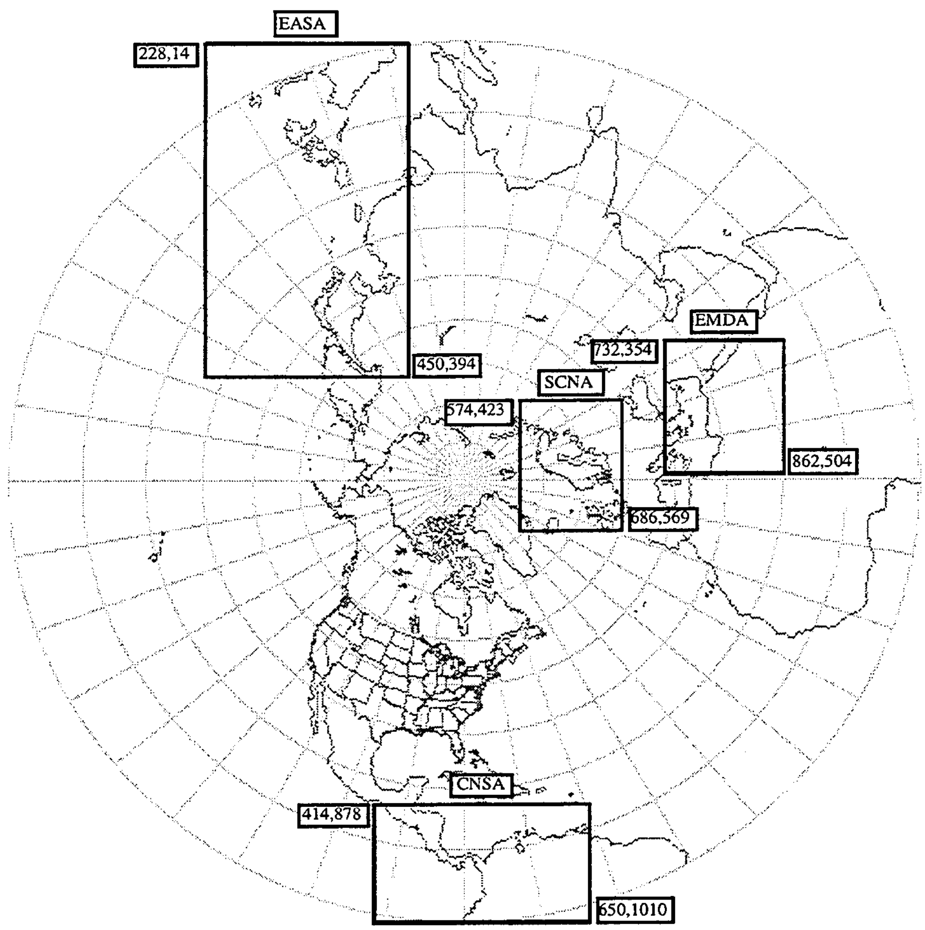
The anticipated scenario for the operational CDFS II cloud model will be two stage. The first will be an event-driven process to analyze raw satellite data through processing layers one and two as they are received at AFGWC. Individual algorithm results will be stored separately in a local database (see Figure 2). The second stage will be a schedule-driven operation wherein the final cloud analysis will be produced through integration of the most timely satellite analyses available in the local database. A global integrated cloud analysis will be produced hourly and stored in a central database on the AFGWC 16th mesh grid. The following parameters will be available: total cloud, number of cloud layers (up to four), fractional amount, type, and height for each layer, and a data quality flag indicating the relative reliability of the analysis for each grid box.

Compared to RTNEPH cloud products, results from the CDFS II cloud model are expected to improve. Use of geostationary data will greatly increase the temporal resolution of the cloud products, particularly in the tropics where polar coverage is poorest and geostationary coverage is best. It is well established that the RTNEPH is deficient in the tropics due largely to strong diurnal variations in cloud amount in this region not captured in the periodic coverage available from polar satellites. This problem is made worse by the timing of the early morning/evening DMSP passes which, when persisted, generally become out of phase with cloud amount trends. Similarly, use of multiple polar satellites (i.e., both civilian and military) will increase temporal coverage for all areas of the globe. Improved cloud detection accuracy from analysis of five-channel AVHRR data, relative to two-channel OLS data, can also be expected. SERCAA algorithms exploit AVHRR multispectral data to improve accuracy for a number of recognized RTNEPH problem situations: low cloud detection, snow/cloud discrimination, classification of optically thin cirrus, over analysis along coastlines and over deserts, and sun glint/cloud discrimination. Increased spectral resolution of OLS data should also result in improved cloud detection and layer discrimination. By moving from 6 to 8 bit pixel resolution, the granularity of IR brightness temperatures improves from 1.9 to 0.5 K. This is particularly significant for detection of low clouds which tend to be thermally indistinct from terrestrial surfaces. In general, analysis of sensor data in raw satellite scan projection should minimize distortion caused by warping and truncation of data which in turn should minimize false cloud features that are artifacts of the remapping process (e.g., over analysis near coastlines caused by inaccurate Earth location of satellite data).

In summary, the CDFS II cloud analysis product is expected to be superior to the currently available RTNEPH analysis. Hourly global analyses will be available at 16th mesh grid resolution vice the current event driven 8th mesh analyses. Introduction of additional satellite sources will greatly improve temporal resolution of the nephanalysis, particularly in the tropics. Use of multispectral AVHRR data should result in improved accuracy in many situations identified as problematic for the RTNEPH. New database techniques in CDFS II are also expected to benefit nephanalysis accuracy through support

for higher spatial and spectral resolution data. Improvements in the cloud analysis retrievals will be of direct benefit to cloud forecast accuracy through better depiction of diurnal trends in cloud cover, more accurate initial conditions and first guess fields for cloud forecast models, and enhanced temporal and spatial resolution.

The DNA data requirements called for delivery of up to eight data sets between September 1994 and March 1995. Each set consists of hourly cloud products and all input data for 10-day continuous periods analyzed from the months of March and July for four geographic locations (Figure 2). The geographic regions were selected by DNA as representative of the kinds of climatic and geographic conditions most stressing to the analysis and forecast models. Initially, another ROI centered over NE Canada was selected; however, due to problems obtaining DMSP data over that region it was dropped in favor of the SCNA region in Figure 2. In addition to the four ROIs identified in the figure, it was decided to also include data from an earlier SERCAA study for three small regions over the Southeast Asian land mass and Western Pacific as an early demonstration set. Ultimately, funding limitations restricted the number of large-ROI data sets produced for DNA to four, two from EASA and one each from CNSA and EMDA. Summaries of data set attributes for each ROI are provided in Tables 5 through 8.



Numbers associated with each ROI indicate 16th mesh grid box bounds.

Figure 2. Regions of Interest For DNA Data Study

Sensor data from four polar orbiting environmental satellites (DMSP F10 and F11, NOAA 11 and 12) and four geostationary (METEOSAT 3 and 4, GOES 7, and GMS 3) were collected to support the DNA program. Each data set covers a 10-day period, normally the last ten days of a month. Various data sources were used to obtain the required satellite sensor data for the selected time periods and ROIs (Table 2). GOES 7 and METEOSAT 3 and 4 data were collected at the PL AIMS facility, GMS data were obtained from the University of Hawaii (through the SeaSpace Corporation), DMSP data were provided by the NOAA National Geophysical Data Center (NGDC), and NOAA data by the NOAA National Climatic Data Center (NCDC). Data from each source were received in different formats and data quality varied widely. It was necessary to develop different data ingest and quality assurance software for each data source. Ingest products included calibrated infrared brightness temperature and visible count data, Earth location information, and sun-satellite geometry information. All ingest data were maintained in the original satellite scan projection of the respective satellite systems.

Table 2. Satellite Data Sources

SATELLITE PLATFORM	DATA SOURCE
DMSP F10 F11	National Geophysical Data Center (NGDC) Boulder, CO
GMS GMS-4	SeaSpace Corporation San Diego, CA
GOES GOES-7	AIMS Direct Readout Ground Station Phillips Laboratory Hanscom AFB, MA
METEOSAT METEOSAT-3 METEOSAT-4	AIMS Direct Readout Ground Station Phillips Laboratory Hanscom AFB, MA
NOAA NOAA-11 NOAA-12	National Climatic Data Center (NCDC) Ashville, NC

Software originally developed to support algorithm testing during the SERCAA program had to be extensively modified to handle the large data sets produced for DNA. Estimates for computer processing time and manpower requirements to analyze data for one ROI using the original SERCAA codes are summarized in Table 3. The table assumes that 80 polar orbiting passes (i.e., 2 per day from 4 satellites) and 240 geostationary scans (i.e., 24 per day from one satellite) will be processed. Estimates for data ingest processing are also included for completeness (i.e., conversion of data received from various sources to the standardized SERCAA DataBase (SDB) format). It is important to realize that SERCAA test programs were developed for the sole purpose of testing the cloud algorithms developed under the program. They were not designed to process large amounts of data in a production sense and, thus, were modified to bring the values in Table 3 down to reasonable levels.

Table 3. Level of Effort Requirements to Process Satellite Data for One 10-Day Period

PROCESSING LEVEL	SATELLITE	MANPOWER (hours)	COMPUTER (wall clock hours)	DISK SPACE (Gbytes)
Level 1 (Ingest)	NOAA	24	90	1.2
	DMSP	120	160	3.0
	GEO	60	40	0.6
Level 2	NOAA	12	20	0.2
	DMSP	48	40	1.5
	GEO	48	20	0.3
Level 3[1]	All	210	660	< 0.1
Level 4	N/A	15	30	< 0.1

[1] Level 3 processing times vary considerably depending on the amount of cloud detected in each ROI, they range from 2-5 hours per ROI.

As previously discussed, the SERCAA cloud analysis algorithms use four levels of processing as summarized in Figure 1. The first level of processing (**Level 1**) consists of ingestion of the data. Tape data are processed through separate ingest programs depending on the data source and format. Modifications to the SERCAA software was required to support the specific DNA test site locations as well as new DMSP and GMS data formats. In addition, modifications to the software was necessary to accommodate multiple orbits of NOAA/AVHRR data.

All Level 1 data are stored in a standard format in the original satellite scan projection. These data are maintained on AIMS through the SERCAA Database (SDB) management software. Level 1 data products consist of separate files for each sensor channel plus two additional files containing Earth location and satellite/solar geometry information. Ingest products are described more completely in Section 2 of Gustafson et al. (1994).

The second level of processing (**Level 2**) employs sensor specific nephanalysis algorithms. Level 1 sensor data from DMSP, NOAA, and the geostationary satellites are processed through separate nephanalysis algorithms as indicated in Figure 1. Each time data from a new satellite pass are ingested, they are analyzed through the appropriate nephanalysis algorithm and results are placed in a Level 2 output file. One output file is generated for each nephanalysis run and nephanalysis results are stored in original satellite scan projection.

The third level of processing (**Level 3**) uses Level 1 and 2 products as input to segment the cloudy regions into vertical cloud layers as well as to classify different cloud types. The process also remaps the data from the individual satellite scan projections to the AFGWC standard polar stereographic map projection (Hoke et al., 1981) at 16^{th} mesh grid resolution. Level 3 products are generated for each 16^{th} mesh grid cell with a maximum of four cloud layers possible for each cell. One Level 3 file is created for each set of Level 1 and 2 products. All Level 1, 2, and 3 products associated with a single satellite pass are related through SDB and are provided on the DNA tapes as a set.

The fourth and last level of processing (**Level 4**) is a clock driven process with one new Level 4 integrated analysis performed each hour. Thus, integration is differentiated from the Level 1, 2, and 3 products that are event-driven. The integration module operates on the most recent Level 3 gridded products available from each satellite source. As was the case with Level 3 products, the Level 4 output files conform to the AFGWC 16th mesh grid structure.

Modifications to the SERCAA software was required in order to process the relatively large volume of data contained in each of the DNA data sets, as compared to the data volume processed during the SERCAA effort. These modifications included the rehosting of the cloud layering and integration algorithms (Level 3 and 4 processing) from the AER computer system, where they were implemented during SERCAA, to the Silicon Graphics, Inc. (SGI) computers on AIMS. This, in turn, required the development of new automated communication software to handle database management between the SGI computers and preexisting VMS-based SDB software. The cloud layering algorithm was implemented using routines from the Image Processing Workbench (IPW) software package. The rehosting and optimization effort for cloud layering included coding the algorithm as stand-alone code with hooks to the appropriate IPW routines.

Hardware upgrades to AIMS were also required in order to improve the efficiency of the satellite data set production process. This hardware was required in order to accommodate the approximately 3 Gbytes of data products that comprise each data set. To achieve this, several key hardware items were purchased by the government for AIMS including two 8mm tape drives, two 1 Gbyte disk drives and 128 Mbytes of additional memory for the SGI computers.

Satellite data sets were collected and processed for regions located in climate specific locations of the world, specified by DNA and listed in Table 4. Three of these regions (HIM, JPN, PAN) are listed as initial data sets. These initial data sets are limited in their coverage area and were provided in order to enable the worldwide cloud forecast model developers with a means of gaining experience with the SERCAA data prior to the delivery of the larger data sets (CNSA, EASA, EMDA). Tables 5 through 8 provide descriptions of each of the individual satellite data sets.

Table 4. DNA Satellite Data Set Regions

REGION NAME	DESCRIPTION
HIM	Himalayas *
JPN	Japan *
PAN	Panama *
CNSA	Central and Northern South America
EASA	Eastern Asia
EMDA	Eastern Mediterranean and Desert North Africa

* Initial Data Set

Table 5. Initial DNA Satellite Data Set Attributes

DATA SET DESCRIPTION	DATA SET ATTRIBUTES HIMALAYAS	DATA SET ATTRIBUTES JAPAN	DATA SET ATTRIBUTES PANAMA
Region Name	HIM	JPN	PAN
Collection Period: Dates Julian Dates	 27-30 May 1993 93147 - 93150	 27-30 May 1993 93147 - 93150	 27-30 May 1993 93147 - 93150
Satellites	DMSP F10 DMSP F11 GMS-4 NOAA-11 NOAA-12	DMSP F10 DMSP F11 GMS-4 NOAA-11 NOAA-12	DMSP F10 DMSP F11 GOES-7 NOAA-11 NOAA-12
(i,j) 16^{th} Mesh Grid	536, 168 - 600, 232	344, 234 - 408, 298	504, 914 - 568, 978
i Range	$536 \leq i \leq 600$	$344 \leq i \leq 408$	$504 \leq i \leq 568$
j Range	$168 \leq j \leq 232$	$234 \leq j \leq 298$	$914 \leq j \leq 978$
Output Size	65 x 65 Grid Cells	65 x 65 Grid Cells	65 x 65 Grid Cells

Table 6. EASA Satellite Data Set Attributes

DATA SET DESCRIPTION	DATA SET ATTRIBUTES	
Region Name	EASA-1	EASA-2
Collection Period: Dates Julian Dates	 22-30 March 1993 93081 - 93089	 22-31 July 1993 93203 - 93212
Satellites	DMSP F10 DMSP F11 GMS-4 NOAA-11 NOAA-12	
(i,j) 16^{th} Mesh Grid	227, 13 - 451, 395	
i Range	$227 \leq i \leq 451$	
j Range	$13 \leq j \leq 395$	
Output Size	225 x 383 Grid Cells	

Table 7. EMDA Satellite Data Set Attributes

DATA SET DESCRIPTION	DATA SET ATTRIBUTES
Region Name	EMDA
Collection Period: Dates Julian Dates	 12-21 March 1994 94071 - 94080
Satellites	DMSP F10 DMSP F11 METEOSAT-4 NOAA-11 NOAA-12
(i,j) 16^{th} Mesh Grid	731, 353 - 863, 505
i Range	$731 \leq i \leq 863$
j Range	$353 \leq j \leq 505$
Output Size	133 x 153 Grid Cells

Table 8. CNSA Satellite Data Set Attributes

DATA SET DESCRIPTION	DATA SET ATTRIBUTES
Region Name	CNSA
Collection Period: Dates Julian Dates	 22-31 March 1994 94081 - 94090
Satellites	DMSP F10 DMSP F11 METEOSAT-3 NOAA-11 NOAA-12
(i,j) 16^{th} Mesh Grid	413, 877 - 651, 1011
i Range	$413 \leq i \leq 651$
j Range	$877 \leq j \leq 1011$
Output Size	239 x 135 Grid Cells

Feedback from users of the satellite data sets identified several problems in the data that were determined to have been caused by program bugs and they were corrected. These problems were: 1) a small percentage of clear pixels in the Level 3 analyses were incorrectly classified as missing and 2) data source information carried in the Level 4 audit trail contained valid entries for data that were no longer contributing to the integrated analysis because they had exceeded the age threshold.

Problems were encountered over the EMDA region with use of the available surface temperature climatology data that affected the accuracy of the Level 2 processing, particularly for METEOSAT data. The diurnal temperature trends observed in the climatological temperatures is out of phase with the clear-scene satellite measurements. There was an approximate six to nine-hour shift in periods of peak cooling and heating between the clear-scene satellite measurements and surface temperature climatology that negatively impacted cloud analysis accuracy. This shift is depicted in Figure 3. Due to this fact, a new reference background temperature database was developed for processing of METEOSAT data using the same technique developed for generation of visible clear-scene reference backgrounds. This result provided good results over the desert background where the problem was most pronounced.

Periodic data dropouts and/or bad radiance values were found on numerous source data tapes that AER received for processing from NGDC. Considerable manual quality control efforts were required to identify and flag the most severely affected data files of which DMSP orbits were identified as the primary source of these bad data. This effort was performed during Level 1 processing so that the most severely affected data files could be eliminated from further processing. Marginally affected orbits were allowed to pass through Level 1 processing with intermediate product and integrated analysis quality checks being performed again during Level 2 and 3 processing. If the quality control check determined that the intermediate product was of questionable quality then the entire orbit was removed from the satellite data set. Thus, any gaps in coverage on the delivered satellite data set tapes are due to either missing or bad data.

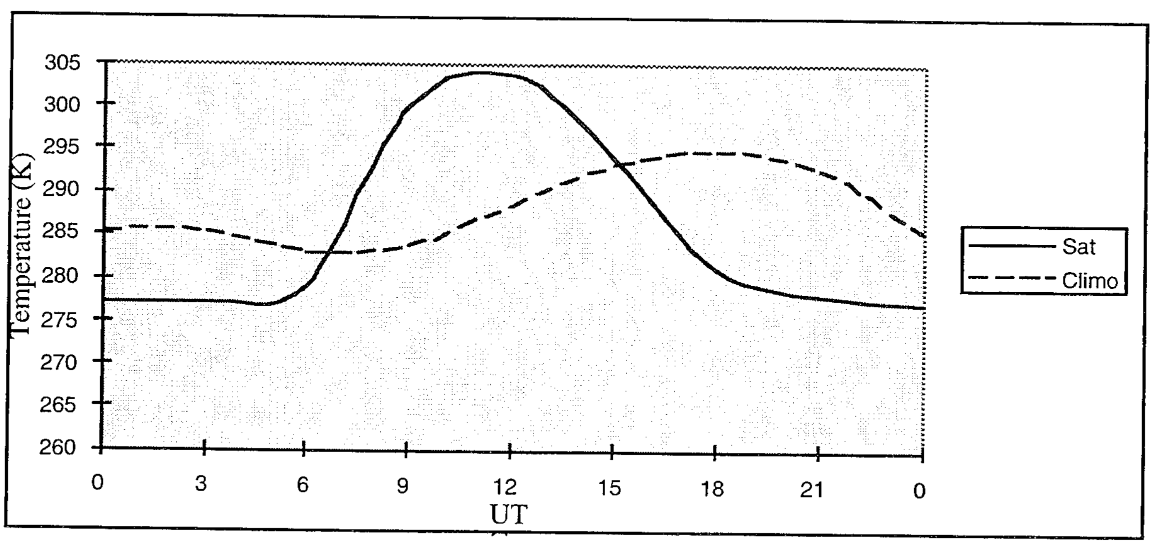


Figure 3. Diurnal Temperature Trends Over the EMDA Region

2.1 SATELLITE DATA SET TAPE ATTRIBUTES

A total of ten 8mm tapes were provided to DNA over the duration of the project. Table 9 provides a listing of these satellite data set tapes and their attributes. For data archiving purposes, all Level 1-3 products associated with a given satellite pass were placed in a single directory and subsequently stored on tape as a single tar file. Thus, the first tape for a collection period contains a series of several hundred tar files; each file contains all Level 1-3 products associated with a single satellite pass. Level 4 files are grouped on the second tape by day. This methodology was not followed for the two data set tapes for the EMDA region. Both of the tapes for the EMDA region contain Level 1-4 products with the each tape covering different intervals of the collection period.

For each set of Level 1-3 products, and for each Level 4 file there is also an SDB Information File associated with it. These files contain descriptive metadata information extracted from the SERCAA Database that describe the relevant attributes of the SERCAA product files.

Table 9. Satellite Data Set Tape Attributes

REGION NAME	COLLECTION PERIOD	TABLE LABEL	CONTENTS
EASA-1	22-30 March 1993	DNA MAR93 ENTRIES	Level 1-3 Products
	93081 - 93089	DNA MAR93 IA	Level 4 Products
HIM, JPN, PAN	27-30 May 1993	SERCAA RE MAY93 1	Level 1 & 2 Products
	93147 - 93150	SERCAA IA MAY93 1	Level 4 Products
EASA-2	22-31 July 1993	DNA JUL93 ENTRIES (RE)	Level 1-3 Products
	93203 - 93212	DNA JUL93 IA	Level 4 Products
EMDA	12-21 March 1994	DNA MAR94 EMD IA/RE 071-078	Level 1-4 Products for first 8 days
	94071 - 94080	DNA MAR94 EMD IA/RE 079-080	Level 1-4 Products for last 2 days
CNSA	22-31 March 1994	DNA MAR94 CNS ENTRIES (RE)	Level 1-3 Products
	94081 - 94090	DNA MAR94 CNS IA	Level 4 Products

3. GOES-8 SATELLITE RECEIVING STATION

On April 13, 1994, the US launched the first in a new series of geosynchronous weather satellites known as GOES-Next. The first satellite of the series, designated GOES-8, was placed in operational status on June 11, 1995 after completing a period of engineering tests and an operational demonstration phase. The next satellite in the GOES-Next series, GOES-9, was launched on May 23, 1995 and was placed in operational status on January 22, 1996. The final satellite of the previous generation of GOES satellites, GOES-7, has been repositioned to a West Coast station and placed on standby mode.

GOES-Next satellites have a different and dramatically more-capable sensor payload than the previous GOES-series as well as a new data format termed GOES Variable (GVAR). For the AIMS system to maintain currency, and to provide the state-of-the-art cloud data sets required by this contract, AER was directed by the government to review alternatives to upgrading the system for GOES-8 and eventually a simultaneous GOES-8 and GOES-9 capability.

We reviewed two possible approaches to achieving a GOES-8 capability. The first relied on an existing system that included an 8-meter tracking antenna, Aydin 1050 subsystem (demodulator, bit sync and frame sync), and an Integral Systems Downlink Interface. The groundstation computer was an Encore 3/67 minicomputer running the MPX-32 realtime operating system. While the antenna had ample G/T to receive the high quality GOES-8 downlink, this approach presented a number of problems. The system was over 20 years old and contains obsolete equipment in both the RF chain and for the antenna mechanical systems. The processing software for the controlling computer was developed in-house and not readily upgraded to GOES-8. Pursuing this approach would require the following:

- extensive software development to either upgrade the GOES-7 software for GOES-8 or to interface it to a turn-key software product (e.g., the SeaSpace system described below)

- uncertain number of engineering hours to work each of the components in the RF chain and make the necessary changes including feed assembly, bit-synchs, and frame-synchs

- severe risk of a major mechanical failure which could not readily be repaired due to parts obsolescence

An alternative approach was to modify an available METEOSAT ground station. This station was receiving METEOSAT-3 data which was stationed over the central Atlantic Ocean. Its data would be made largely redundant by GOES-8 and thus will no longer be needed when GOES-8 capability is available. The METEOSAT ground station was manufactured by SeaSpace and is fully supported by them - i.e., it has none of the parts obsolescence problems of the first alternative. The following were required to modify the system:

- a new LNA, feed

- new down-converter will be provided

- new bit synchronizer
- reprogram the METEOSAT frame synchronizer
- upgrade the METEOSAT Sparc IPX workstation
- move the METEOSAT antenna (which was located at AER's Cambridge facilities) and install on the roof of the Phillips Laboratory (GPAB)
- complete integration and testing at Hanscom AFB

A quote was received from SeaSpace to perform the above tasks (except for the move and reinstallation of the antenna for which a separate quote was obtained). AER reviewed these alternatives with the government, who concurred with our recommendation of proceeding with the second approach - modification of the METEOSAT ground station. In July of 1995, the modifications were completed and the GOES-8 groundstation was installed at PL. The purchase of the GOES system from SeaSpace Corporation maintained a consistent base of weather satellite processing systems, allowing computer hardware and software to be easily integrated into the existing computer network and maintaining consistency in vendor-supplied software that scientists and support personnel are already familiar with to ingest and process weather satellite data.

The GOES-8 system technical specifications are as follows:

ACQUISITION ELECTRONICS:
- Paraclipse 12' antenna, LNA and downconverter
- SeaSpace HR100 receiver/bit synchronizer
- SeaSpace S-BUS frame synchronizer

COMPUTER HARDWARE:
- SUN Sparcstation 20 with:
  64 MB system memory
  1 GB system disk
  4 GB user disk
  Two 1 GB pass disks
  CD-ROM drive
  8mm stacker

COMPUTER SOFTWARE:
- Solaris 2.4 operating system
- Terascan 2.6 satellite ingest and processing software

The following summarizes the chronology of the GOES-8 ground station installation/integration:

- 3/95: final purchase orders placed
- 4/95: Sparc IPX memory and disk upgrades made; Sparc shipped to SeaSpace

for system integration

- 5/95: METEOSAT antenna moved from Cambridge to PL/Hanscom
- 7/95: SeaSpace completes integration and testing at their facilities and ships equipment
- 7/95: system installed and tested at PL/Hanscom; basic operating capability demonstrated
- 8/95: regular archiving of GOES-8 data initiated

Among the future tasks relating to GOES-Next are:

- Simultaneous GOES-8/-9 capability (will require entire new system as there is no spare hardware available)
- Develop software for ingest/processing of sounder data
- Acquire a more powerful computer for ingesting GOES data

The general imaging operation of the GOES-8 satellite, located at 75 degrees west longitude, is on half-hour boundaries; once at 15 minutes past the hour and again at 45 minutes past the hour. In the routine mode of operation, GOES-8 provides frequent full earth, northern hemisphere, continental US (CONUS) and small sector southern hemisphere frames. Table 10 shows the geographic extent of the frames used in the GOES-8 operations while Figure 4 illustrates how these frames fit into the operational schedule. Note the five half-hourly cycles comprised of the extended northern hemisphere, CONUS, and southern hemisphere (small sector) frames that divide the three-hourly full disk cycle. Similar information for GOES-8 sounder operations is shown in Table 11 and the lower half of Figure 4.

Table 10. Geographical Definitions for GOES-8 Imaging Frames

FRAME NAME	BOUNDARIES (Lat/Long as viewed from spacecraft)			
	North	South	West	East
Full Disk	Earth Edge			
Full Disk - Abbreviated	90°N	51°S	Earth Edge	
Full Disk	90°N	23°S	Earth Edge	
Northern Hemisphere	66°N	2°N	117°W	36°W
Northern Hemisphere-Extended	66°N	20°S	117°W	36°W
Southern Hemisphere-South	20°S	50°S	117°W	36°W
Continental US (CONUS)	61°N	14°N	111°W	62°W
Southern Hemisphere-Small Sector	0°	15°S	114°W	81°W

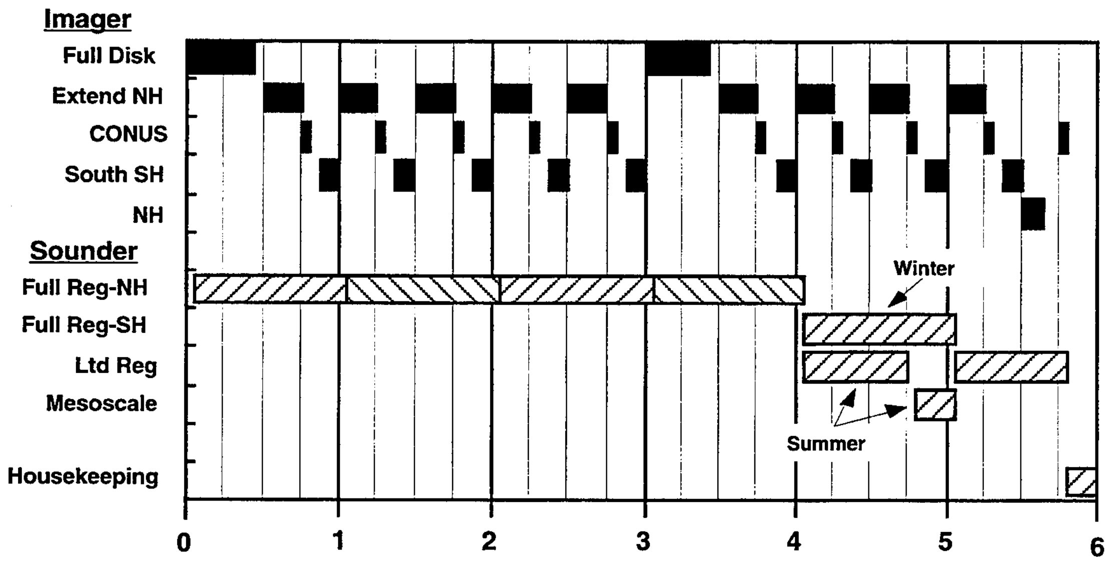


Figure 4. Routine Mode Operational Schedule for GOES-8 Imaging and Sounding

Table 11. Geographical Definitions for GOES-8 Sounding Frames

FRAME NAME	BOUNDARIES (Lat/Long as viewed from spacecraft)			
	North	South	West	East
Full Regional-NH	51° N	23° N	121° W	64° W
Limited Regional-NH	50° N	26° N	120° W	66° W
Full Regional-SH	22° N	50° S	121° W	64° W
Meso-Tropics 4	23° N	11° N	115° W	93° W

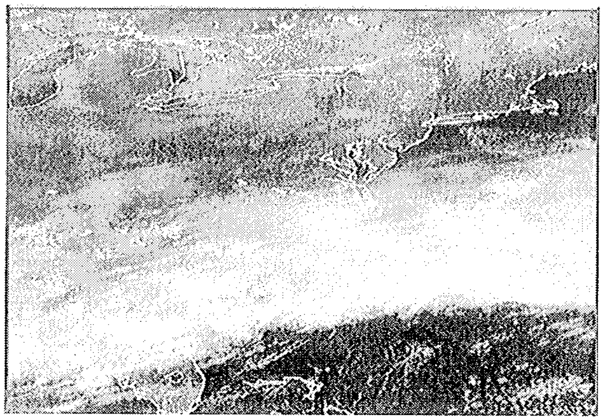
The GOES-8 system has been configured so that SeaSpace realtime acquisition software runs continuously, allowing all frames broadcast by the satellite to be detected and processed. This software stores all data from the telemetry stream to one of four partitions configured over two disks (known as pass disks). Data stored on these partitions are highly volatile since they are used in a round-robin fashion, driven by each satellite frame detected in the broadcast stream. Typically, once a frame has been completely ingested, post-processing software is activated to extract sensor and ancillary data, format the data, and archive it to 8mm tape. An alternative to processing data from disk is to process the data in realtime as it is received. This is a capability new to SeaSpace systems that allows one or more user-generated scripts to be executed during data acquisition. A script will contain SeaSpace Terascan commands that allow the selection, spatial and spectral resolution, and engineering units of the data. Scripts can also contain UNIX commands for performing post-processing activities such as data compression and data archive. The advantage of processing data in realtime (versus dumping the raw data to partitions) is once ingest is complete for a frame or a user-

defined subset, the data is available for use. With data on the pass disks, the user must wait for the entire frame to be ingested and then must run software to reformat the data from the raw stream. As an example, a user-defined subset describing CONUS could be extracted from an extended northern hemisphere frame and be made available for use 7 minutes after the start of the frame. The same subset wouldn't be available until 17 minutes after the start of frame if processed from a pass disk partition.

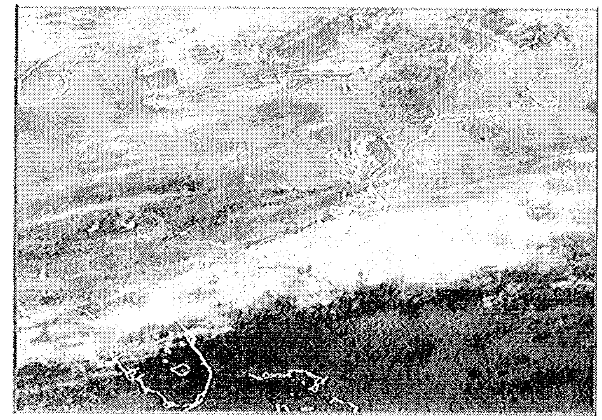
The GOES-8 imaging instrument components include eight visible channel detectors linearly aligned in the north-south direction that are sampled simultaneously and digitized as 10-bit words to provide imagery with a nominal resolution of 1 km square at nadir. Seven thermal detectors of two different sizes sense infrared radiation in four spectral channels. Three of the channels, 3.9, 10.7 and 12.0 µm central wavelength, employ small detector pairs that are simultaneously sampled and digitized as 10-bit words to provide imagery with a nominal resolution of 4 km square at nadir. The 6.75 µm channel employs the large detector, also providing 10-bit data with a nominal resolution of 8 km square at nadir. All IR detectors have redundant counterparts. Figure 5 provides examples of imagery from five of the GOES-8 sensor channels. The region shown is over the Eastern US.

A secondary issue concerns the high data rates and volumes of GOES-8. The Sparc IPX may not be suitable for extensive exploratory data analysis, display, calibration, geolocation, and other processing of the GOES imagery. An interim solution is to perform these functions on the DMSP Sparc workstation. This approach incurs no additional costs, but it is not clear how acceptable it will be overall. A preferred long term solution is to acquire another Terascan License for one of the SGI Indigo workstations. This would enable off-line post-processing and value added processing to occur on a high end workstation, well-suited to the large data volumes of GOES-8. It was decided to defer purchase of a Terascan License for the SGI Indigo workstation until a later time.

Since August 1995, half-hourly northern hemisphere frames containing data from all five channels have been archived to 8mm tape as UNIX tar sets. IR data are stored full resolution as brightness temperatures while visible data are subsampled to the lower 4 km IR resolution and stored as scaled percent albedo. Data quality has generally been excellent with little or no line dropouts.



Channel 1
Visible (0.55 - 0.75 μm)



Channel 2
Infrared (3.80 - 4.00 μm)

Channel 3
Infrared (6.50 - 7.00 μm)

Channel 4
Infrared (10.20 - 11.20 μm)

Channel 5
Infrared (11.50 - 12.50 μm)

Figure 5. Examples of GOES-8 Sensor Channel Images

4. SPACE-BASED INFRARED SYSTEM (SBIRS)

The SBIRS project requires the completion of two major tasks. These tasks are to 1) perform an evaluation of alternative cloud analysis algorithms in order to develop a climatology of cloud parameters for a single region and season combination based on the algorithm set selected, and 2) provide cloud products (based on the optimized algorithms selected in task one) for specific regions and times coinciding with observations made by DoD infrared sensing missions that are described below. Currently, task one is in progress and task two has yet to be commenced.

As mentioned, the objective of the first SBIRS task is to develop a climatology of cloud parameters based on SERCAA cloud analysis algorithms. Analyzed output frequency is 3-6 hours (TBD) over a 20-25 km grid. The required parameters are layer cloud fraction and altitude. This task required the collection of NOAA/AVHRR HRPT and GOES-8 GVAR imager data, received by the Phillips Laboratory direct broadcast satellite ground stations, over a designated study area. Since all satellite data collection was performed using Phillips Laboratory GOES and NOAA/POES ground stations, selection of the study area was necessarily limited by line-of-site constraints for the NOAA polar-orbiting data and the data archive capabilities for the GOES data. The study area selected includes the Northeast United States, adjacent ocean, and extends far enough west to include Madison, WI. Satellite data were collected over this region during the period from August through September, 1995.

Evaluation of the SERCAA Phase I (Gustafson et al., 1994) AVHRR and Geostationary algorithms as well as the SERCAA Phase II (d'Entremont et al., 1996) AVHRR and GOES-8 algorithms are being performed in conjunction with analysis of the satellite data collected over the study areas. Reference cloud base and top altitudes are derived with the use of near-coincident ground-based TPQ-11 radar located at Hanscom AFB and lidar measurements collected at the University of Wisconsin, located in Madison, WI. Cloud altitudes are defined by applying thresholds to these radar and lidar data. Comparisons are then made between the satellite-derived values using both versions of SERCAA algorithms to the reference radar/lidar altitudes. Radar/lidar time-series measurements are compared to areal satellite-retrieved values by converting from (radar/lidar) time coordinates to (satellite) spatial coordinates using a trajectory field based on cirrus altitude winds over radar/lidar site. Comparisons are performed out to ±30 minutes from valid time of satellite data. The comparisons are best done graphically, but standard comparison statistics can also be generated (e.g., mean, standard deviation).

Modifications were made to the SERCAA Phase I geostationary algorithm to exploit the additional full-time GOES-8 imager channels (3.9 and split LWIR) that are now available. Due to the similar channel characteristics between GOES-8 and AVHRR, these modifications were primarily a programming effort. Testing was performed to verify that the changes made to the algorithm provided the results that were expected based on results of the already established and tested SERCAA Phase I AVHRR algorithms. SERCAA Phase I algorithms, for all platforms, were also modified to produce fewer number of cloud layers with greater temporal consistency as compared to SERCAA Phase I results. This modification was made in accordance with feed-back from DNA.

The objective of the second task is to produce analyzed cloud products using the cloud algorithms selected in Task 1 to support specialized SBIRS data collections. Two satellites known as the Midcourse Space Experiment (MSX) and the Miniature Seeker

Technology Integration (MSTI) are scheduled for launch in Spring 1996. These, along with a high-altitude aircraft known as ARES, will be making a series of measurements over various locations around the globe during 1996. Corresponding environmental satellite data will also be collected and provided to AER for cloud analysis processing. Some missions will occur within the Phillips Laboratory satellite coverage area but is assumed that most will not. For those areas that are located outside of the PL coverage area, environmental satellite data will be provided by MIT Lincoln Laboratories through their McIdas tap. Polar-orbiting data will be collected from the orbits intersecting the region of interest that occur immediately before and after the mission time. It is assumed that there may be a difference of up to four hours between the mission time and the time of both polar passes. Geostationary data will be collected for the two closest time periods occurring before and after the mission time plus the time period immediately preceding the first time period (SERCAA geostationary algorithms require data from at least two consecutive times). Approximately 100 mission support cases are expected to be analyzed.

4.1 CIRRUS-PROPERTY RETRIEVAL ALGORITHMS

In addition to cloud amount and altitude information, effective emissivity and optical depth statistics for cirrus using SERCAA Phase II algorithms will also be derived from AVHRR and GOES-8 data. Current versions of these algorithms operate only on nighttime data due to contamination of 3.7 μm measurements by reflected solar during the daytime. Investigations are being performed to determine if it is possible for the algorithm to operate under daytime conditions and also to incorporate them into the Phase I algorithms. Required changes to accommodate daytime data should probably be assumed to be less accurate than the nighttime retrievals. Incorporation into Phase I algorithms should be straightforward except for automating identification of required cloud-free background temperatures.

Cirrus is recognized as one of the most poorly quantified of all clouds. Its altitudes are difficult to specify, because it typically consists of ice particles distributed over a considerable vertical extent, and its optical properties and microphysics are complex. In addition to the wide variability in properties common for other cloud types, cirrus clouds have the distinct complexity of transmissivity values t that span the entire possible domain $0< t <1$. Other uncertainties in satellite-retrieved cirrus attributes include thin cirrus fraction, altitude, and thickness because the measured cirrus signal is affected by both cloud and the ground below. A more accurate determination of cirrus attributes is needed on both global and local spatial scales. Schiffer and Rossow (1983) specify the ISCCP goal of 30-day average cloud fraction to an accuracy of $\pm 30\%$ for global total cloud; required cirrus accuracies are specified to be $\pm 5\%$ for fraction, and ± 1 km for altitude.

The most extensive cirrus climatologies to date are those compiled at the University of Wisconsin since the mid-1980s using the CO_2 slicing technique. CO_2 slicing is used with NOAA geostationary and polar orbiting satellite sounder channel data to retrieve cirrus altitude and effective emissivity (Wylie and Menzel, 1989). Earlier cirrus climatologies (1979 - 1981) include those obtained using SAGE limb sounder data (Woodbury and McCormick, 1983). A solar disc extinction analysis technique was used to determine the presence of high-altitude cirrus at a spatial resolution of 100 km^2. In the early 1970s a limited amount of Nimbus-4 IRIS infrared data were analyzed to detect the presence of cirrus over ocean between $\pm 50^\circ$ latitude (Barton, 1983).

The transmissive nature of cirrus clouds turns out to be its most important (in a climate sense) and elusive (in a retrieval sense) attribute to specify. If the semi-transparent nature of cirrus clouds is not accounted for, its altitude is consistently underestimated when using passive infrared brightness temperature data. Although it is generally agreed that cirrus has a net warming effect on climate, determination of the magnitude of this effect depends critically on the accurate specification of cirrus radiative and spatial attributes. For example, in the case of very thin ("sub-visual") cirrus, ice particles have a more significant interaction effect with incident solar and upwelling thermal radiation than does upper tropospheric water vapor (Smith et al., 1990).

4.2 OBSERVATION OF CIRRUS FROM SATELLITE

There are many sources of passive satellite data that can be used to detect and analyze cirrus attributes. Among the earliest are visible and infrared data of the 1960s from the TIROS series of polar orbiting satellites, augmented in the early 1970s by geostationary GOES data. Current GOES VISSR Atmospheric Sounder (VAS) channels useful for detection of cirrus include 3.9, 6.7, 11.2, 12.7, and CO_2 13.3 - 14.5 μm spectral bands. The 3.9 and 11.2 - 12.7 μm channels will be discussed in detail shortly; the 6.7 μm water vapor band has proven useful for detection of very thin cirrus over warm backgrounds such as deserts and tropical oceans.

More recent TIROS sensors include the Advanced Very High Resolution Radiometer (AVHRR), a five-channel passive radiometer with detectors that measure upwelling visible (0.63 μm), near-infrared (NIR, 0.86 μm), middle wavelength IR (MWIR, 3.7 μm), and split longwave IR (LWIR, 10.7 and 11.8 μm) energy both day and night. The sounder instruments collectively known as TOVS (TIROS Operational Vertical Sounder) also collect data in the wings of the 15 mm CO_2 absorption band that are useful for detection of thin cirrus and specification of their height. There are also very high spatial resolution (500 m) Defense Meteorological Satellite Program data available in visible/NIR (0.4 - 1.1 μm) and LWIR (10 - 12 μm) bands that are helpful in ascertaining the small-scale spatial attributes of cirrus.

4.2.1 Passive Infrared Physics of Cirrus Cloud Signatures

The upwelling spectral thermal radiance I_{obs} measured by a downward pointing radiometer for a field-of-view completely filled by a non-reflective, thin cirrus cloud is

$$I_{obs} = (1 - \varepsilon) I_{sfc} + \varepsilon I_{cld} , \tag{1}$$

where ε is the bulk cirrus emissivity, I_{sfc} is the upwelling radiance emitted by the underlying surface and clear atmosphere, and I_{cld} includes the cirrus blackbody radiance plus the radiance emitted by the atmosphere above the cloud. In practice, the non-reflective nature of cirrus clouds at thermal infrared wavelengths is assumed. This is a reasonable assumption not only because the cirrus bulk reflectivity is low, but also because there is only minor downwelling thermal emission incident on the top of cirrus clouds to be reflected back to space.

In theory, the specification of I_{sfc} in Eq. (1) requires information on many of the physical properties of the atmosphere and surface that underlies the cirrus cloud: the temperature T_{sfc} and atmospheric transmittance τ (for water vapor attenuation) are two of the more important attributes. As discussed later, I_{sfc} is specified using nearby measurements of cirrus-free pixels.

The two unknowns of interest in Eq. (1) are the cirrus bulk emissivity ε and the cirrus Planck blackbody emission I_{cld}, which is a known function of the cirrus effective temperature T_{cld}. In contrast, there is only one known in Eq. (1), namely the radiance I_{obs}. In order to specify these two unknowns, additional measurement information is needed. This is achieved first by considering Eq. (1) for simultaneous radiance measurements at two different infrared wavelengths.

For purposes of discussion, assume that the radiance data are being measured by the AVHRR MWIR Channel 3 and LWIR Channel 4 sensors. The two cirrus radiance equations are then

$$I_{obs,3} = (1 - \varepsilon_3) I_{sfc,3} + \varepsilon_3 I_{cld,3} \tag{2a}$$

and

$$I_{obs,4} = (1 - \varepsilon_4) I_{sfc,4} + \varepsilon_4 I_{cld,4} , \tag{2b}$$

where the "3" and "4" subscripts denote the 3.7 and 10.7 μm AVHRR Channels 3 and 4 radiances, respectively. Eqs. (2a) and (2b) are two equations, but with the second equation a third unknown ε_4 has been introduced.

A third equation is needed that contains no new variables and that relates at least two of the three unknowns already established. This is done by assuming a relationship between the cirrus bulk optical depths δ_3 and δ_4 as follows. First, radiative transfer calculations are available that compute bulk cirrus optical depth as a function of wavelength and hexagonal ice particle size for varying cirrus cloud thicknesses (Takano and Liou, 1989; Hunt, 1973). Once computed, a simple linear regression between corresponding pairs of the two optical depths is performed to obtain a relationship of the form

$$\delta_3 = m\, \delta_4 + b , \tag{3}$$

where m and b are the regression slope and intercept, respectively. The slope m is nonzero; the intercept b, however, turns out to be very close to zero since the two optical depths are close to each other for optically very thin cirrus. Thus, Eq. (3) is generally of the form

$$\delta_3 = m\, \delta_4 . \tag{4}$$

Considering the radiative properties of cirrus within a satellite field-of-view in a bulk sense, the cirrus cloud optical depth δ is related to the cirrus transmissivity t by the relation

$$\delta = -\ln t , \tag{5a}$$

so that Eq. (4) can be written

$$\ln t_3 = m \ln t_4 . \tag{5b}$$

Since it is assumed that the cirrus cloud is non-reflective, $\varepsilon + t = 1$ so that

$$\ln (1 - \varepsilon_3) = m \ln (1 - \varepsilon_4) . \tag{5c}$$

Finally, solving for ε_3 in terms of ε_4 yields

$$\varepsilon_3 = 1 - (1 - \varepsilon_4)^m . \tag{6}$$

This is the third of the three-set equation, so that there are now three equations (2a), (2b), and (6) in three unknowns ε_3, ε_4, and T_{cld}. The three measurements consist of the linear regression slope m, and the satellite-measured radiances $I_{obs,3}$ and $I_{obs,4}$. This three-equation system forms the basis for most cirrus retrieval techniques that analyze multispectral infrared satellite radiances.

In practice, nearby cirrus-free pixels are used to obtain accurate estimates of $I_{sfc,3}$ and $I_{sfc,4}$. Atmospheric emission above the cirrus is neglected. Cirrus reflectivity is neglected as well. As previously mentioned, this does not introduce into the retrieval process a major source of error at night, but during the daytime incident solar radiances at the shorter 3.7 μm wavelengths noticeably affect cirrus radiance measurements. The daytime problem is a challenging one. Although the reflectivities of cirrus are relatively small, incoming 3.7 μm solar radiation is strong enough so that measured radiances are solar-contaminated. Subsequently, Eq. (2a) is no longer accurate and the retrieval process becomes considerably more complex because of it. Thus, the use of Eqs. (2a), (2b), and (6) are presently restricted to nighttime scenes when there is no incident solar energy being reflected back to space by either the cirrus cloud itself or the underlying background. Research is ongoing to separate out the solar component in the 3.7 μm daytime data (Ou et al., 1993).

Another major constraint is that in assigning a single cloud temperature T_{cld} to the cirrus, it is assumed that the cloud is a thin sheet that lies precisely at one atmospheric level. Clearly this is not the case. Lidar backscatter returns from cirrus clouds consistently show their complex structure on both horizontal and vertical scales. In midlatitudes their altitudes range from 6 to 13 km, and their thicknesses anywhere from 1 to 5 km and, on occasion, even higher. Thus, the assignment of a single cirrus temperature is a gross one which, in the case of multispectral infrared radiance retrievals, results in the assignment of one effective cirrus cloud altitude. However, the severity of this constraint affects only the cirrus altitude determination. Its effects on the bulk optical properties of the cloud are far less detrimental. Nonetheless, it is important to remember that current satellite-retrieved cirrus altitudes are not an accurate assessment of the true levels at which the cirrus lie, but rather are only correct in a radiatively bulk, energy-average sense. For this reason, the cirrus temperature T_{cld} and corresponding altitude z_{cld} are labeled as "effective" properties, since they afford little inference on the detailed vertical structure of the cloud.

In practice, Eqs. (2) and (6) can also be used with AVHRR channels 3 and 5. Thus, for every triplet of AVHRR infrared satellite radiance measurements $I_{obs,3}$, $I_{obs,4}$, and $I_{obs,5}$ it is possible to retrieve at sensor resolution the 3.7, 10.7, and 11.8 μm cirrus bulk emissivities ε_3, ε_4, ε_5 and optical depths δ_3, δ_4, δ_5 along with cirrus effective temperature (altitude) T_{cld} (z_{cld}).

4.2.2 Multispectral Infrared Cirrus Brightness Temperature Differences

There are three main reasons why cirrus detection using multispectral infrared measurements is successful at night. The most dominant effect has to do with the nature of the dependence of the Planck function on temperature at the three AVHRR IR wavelengths. To illustrate this, consider the simple example in which a pixel contains a cirrus cloud of temperature 230 K and emissivity 0.5, and under which lies a background surface of 280 K. (For now, ignore atmospheric effects and the fact that the cirrus emissivities are not constant from one channel to the next. These issues are discussed later.) According to Eq. (1), in this situation some of the measured upwelling radiance in

each channel will originate from the cold cloud at 230 K and the rest from the underlying warm surface at 280 K. Using the Planck function, the proportions of the total radiances (half cloud and half background) that are measured by the satellite can be computed; results are plotted in Figure 6. Note that proportionally less energy comes from the warmer part of the scene as wavelength increases; this effect is due solely to the exponential dependence of the Planck function on temperature at the three wavelengths. The resulting brightness temperatures are also plotted in Figure 6; note the high, positive differences $T_3 - T_4 = 8.4$ K ($\Delta T_{3,4}$) and $T_3 - T_5 = 8.9$ K ($\Delta T_{3,5}$). Such brightness temperature differences are consistently and distinctively characteristic of thin cirrus clouds in nighttime AVHRR data.

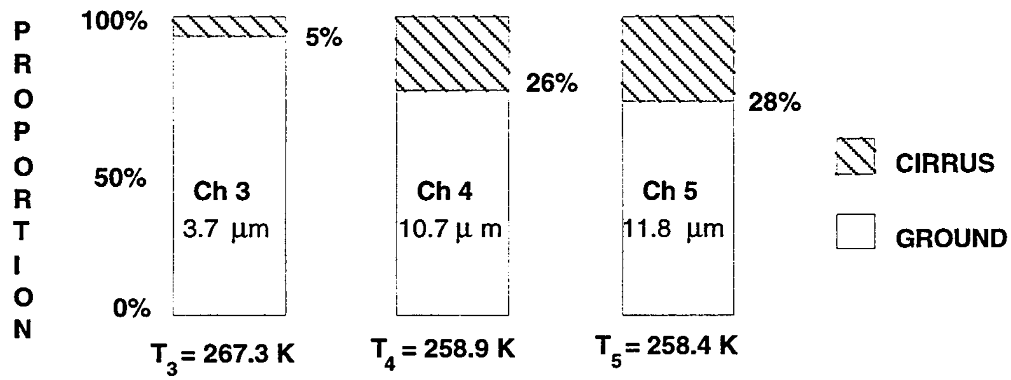


Figure 6. Cirrus Brightness Temperature Difference Plots

The second effect is that of the varying emissivities of ice particle clouds themselves among the three wavelengths. In general, cirrus emissivities increase with increasing wavelength for the three AVHRR infrared channels. This means that, on the basis of emissivity alone, increasingly more of the upwelling radiant energy in a cirrus-filled pixel comes from the colder cloud. This is an analogous but weaker effect to that of the Planck function in that it amplifies the brightness temperature differences which comprise the thin cirrus signature.

Finally, a third effect that causes brightness temperatures for cirrus pixels to decrease with increasing wavelength is that of varying atmospheric water vapor attenuation. In general, atmospheric water vapor attenuation is stronger at longer wavelengths. This operates in the same sense as do the previous two, increasing the difference between channel 3 and channels 4 or 5 brightness temperatures, although typically it is the weakest effect of the three.

T_4-T_5 differences for ice particle clouds are also due to the same reasons; however, the dominant effect in this case is that of changing emissivity. This is mostly due to that fact that channels 4 and 5 are spectrally too close to one another for the Planck temperature dependence differences to manifest themselves as significantly as they do for 3.7 µm. However, even though the differences tend to be smaller, they are very consistent for thin cirrus, especially very thin cirrus.

Figure 7 plots modeled brightness temperature differences as a function of cirrus emissivity, taking into account all three of the effects discussed. In summary, these are: 1) the stronger dependence of the Planck function on temperature at 3.7 µm wavelengths, 2) the differences in emissivities at each of the three AVHRR IR wavelengths as described by Eq. (6), and 3) stronger atmospheric water vapor attenuation in the longer

wavelength regions (i.e., 10.7 and 11.8 μm). Note that interchannel differences are highest for cirrus clouds with bulk emissivities of approximately 0.8 to 0.9, and with higher effective cloud altitudes of 10 to 13 km. The atmospheric temperature and water vapor profiles used in generating the model results of Figure 7 were measured in the presence of thin cirrus clouds during the 1986 FIRE experiment in Wisconsin (d'Entremont et al., 1990).

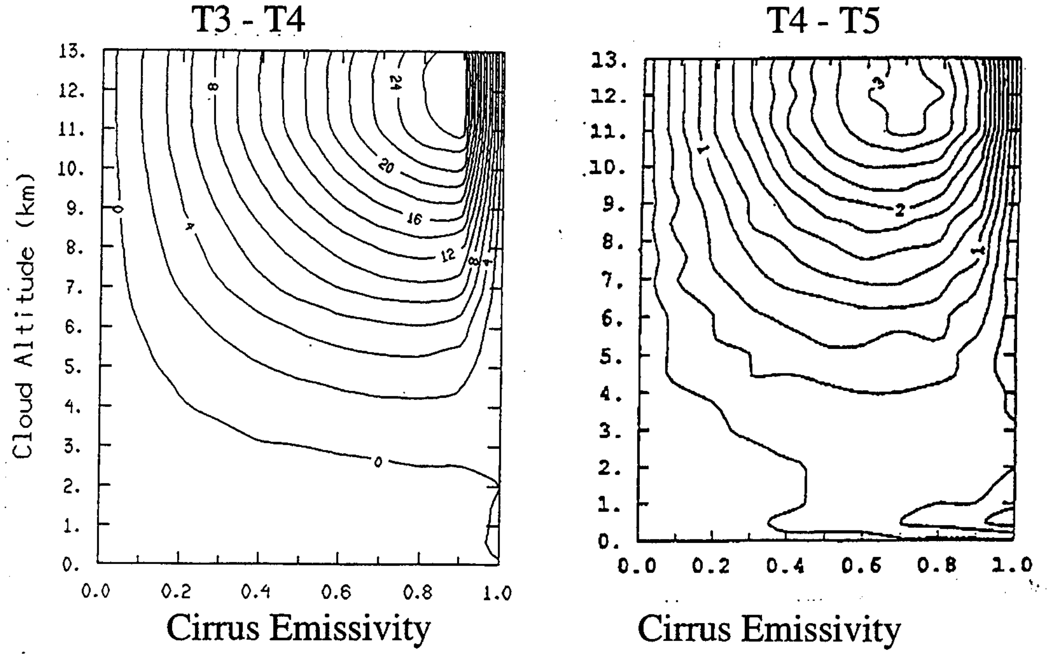


Figure 7. MWIR-LWIR Brightness Temperature Differences

4.2.3 Sensitivity of AVHRR Cirrus Emissivity Retrievals

With so many physical processes contributing to the successful detection of cirrus clouds using multispectral infrared satellite data, it is important to be aware of the extent to which cirrus retrievals are sensitive to accurate radiometric measurements. Figure 8 plots the sensitivity in retrieved bulk emissivity for cirrus clouds of 0.9 (thick), 0.5 (thin), and 0.2 (very thin) emissivities, as a function of radiance measurement error (expressed in the plots in terms of brightness temperature difference ΔT). Positive ΔTs indicate brightness temperature (radiance) measurements that are higher than they should be, and vice versa. Note that for cirrus clouds with high emissivities, the effect of measurement error on retrieved emissivity is not harmful. This makes intuitive and physical sense because when cirrus emissivities are high, most of the radiance measurement comes from the cloud itself, and not the background. However, the effect of error worsens for thinner cirrus clouds. For emissivities of 0.5, a measurements error of ±3 K in both brightness temperature measurements can result in emissivity errors of 0.2 or greater. The effect of measurement error on very thin cirrus clouds is the most severe. For channel 3 measurement errors of -3 K, retrieved emissivity errors of greater than +0.6 result. This means that a cirrus cloud of 0.2 emissivity appears to have an emissivity of greater than 0.8 when the channel 3 brightness temperature measurement is 3 K too low; this is substantial. Although channels 4 and 5 radiance measurements are typically accurate for cirrus clouds, channel 3 radiance measurement errors corresponding to ±3 K in brightness temperature are not uncommon due to sensor noise. It is important to realize this

sensitivity when quantitatively analyzing AVHRR IR data for cirrus clouds. Because of the noise in channel 3, and because of the added complexity involved in processing solar-contaminated daytime 3.7 μm data, increased attention is again being placed on retrieval of cirrus attributes using channels 4 and 5 data only (Parol et al., 1991; Smith et al., 1990).

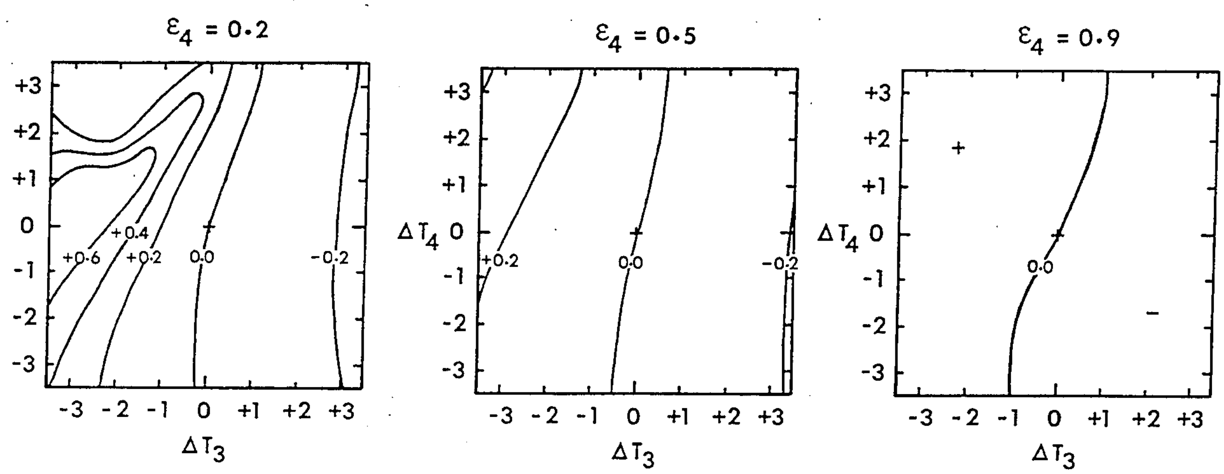


Figure 8. Retrieved Emissivity Sensitivity to Measurement Error, Noise

4.2.4 The CO_2 Slicing Technique

In recent years, the CO_2 slicing technique has been extensively applied to obtain cirrus emissivity and altitude statistics on a global scale (Wylie, Menzel, and Woolf, 1991; Menzel, Wylie, and Strabala, 1992). The CO_2 technique estimates cirrus altitude and the product of cloud fraction (N) and emissivity (ε), called "effective emissivity (Nε)." This technique uses satellite-measured radiances from the NOAA High-Resolution Infrared Radiation Sounder (HIRS) CO_2 channels in the 13.4 - 14.2 μm spectral range. It is capable of detecting the presence of transmissive cirrus clouds at levels in the upper troposphere above where the CO_2 channels' weighting functions peak. The differential absorption characteristics within the HIRS spectral bands of the radiation passing through clouds allows the CO_2 slicing method to detect transmissive clouds and specify their height. This is one of the strongest attributes of the technique. Once the presence and effective altitude of transmissive cloud is established, the HIRS Channel 8 (11.1 μm) IR brightness temperature measurement T_8 is used along with an estimate of the surface skin temperature T_{sfc} to compute effective emissivity. If the CO_2 channel radiances fail to reliably detect the presence of transmissive cloud, an emissivity of 1.00 is assumed and T_8 is compared to the temperature profile to assign a blackbody, opaque cloud top altitude. Cloud climatologies produced thus far using the CO_2 slicing method have been extensive, focusing on determining the geographical, seasonal, and diurnal changes of cloud cover (Wylie and Menzel, 1989). CO_2 slicing cloud analyses have been compared with lidar measurements and NWS ground-based cloud reports (Wylie and Menzel, 1989). The first-order radiative characteristics of transmissive cirrus clouds are becoming better understood on a global scale using the CO_2 slicing climatologies. A detailed description of the CO_2 slicing technique is given by Wylie and Menzel (1989).

4.3 TESTS AND RESULTS

Nighttime AVHRR data for Channels 3, 4, and 5 were obtained over New England at ~2337 UTC on 16 September 1995, a time when surface-based radar observations of thin cirrus were available at Hanscom AFB as a part of the SBIRS field observations. The Hanscom site was equipped with the ground-based active radar system. The TPQ-11 is a 35GHz upward-pointing radar that provides useful observations of cirrus cloud base and top against which satellite-based retrievals can be directly compared. AVHRR Local Area Coverage (LAC) data were used. Each LAC pixel has a resolution of approximately 1 km at nadir. The AVHRR image sample over Hanscom is close to satellite nadir, and was selected for the following reasons: 1) there were abundant thin cirrus in the southern New England area; 2) atmospheric sounding information from nearby RAOB stations was readily available, valid for the time and location of the thin cirrus observations, and 3) surface-based radar observations of the cirrus were collected for a period before, during, and after the NOAA satellite overpass time. The radar observations serve as a best source of ground truth for verifications of calculated z_{cld}.

Subsequent cloud analysis results are then compared to the cirrus radar observations to help verify (or at least substantiate) the calculations of z_{cld}. It is not possible to verify the cirrus emissivity estimates except to say that if the effective cirrus altitudes z_{cld} are reasonable, then the emissivities are also likely to be reasonable since these two parameters are coupled in the physical retrieval model.

Figure 9 plots results of the enhanced cirrus retrieval over Hanscom AFB using AVHRR infrared Channels 3 and 4. The retrieved cirrus effective altitudes agree well with the radar observations of cloud base and top, lending confidence that the transmissive characteristics of the thinner cirrus are being properly accounted for in the model.

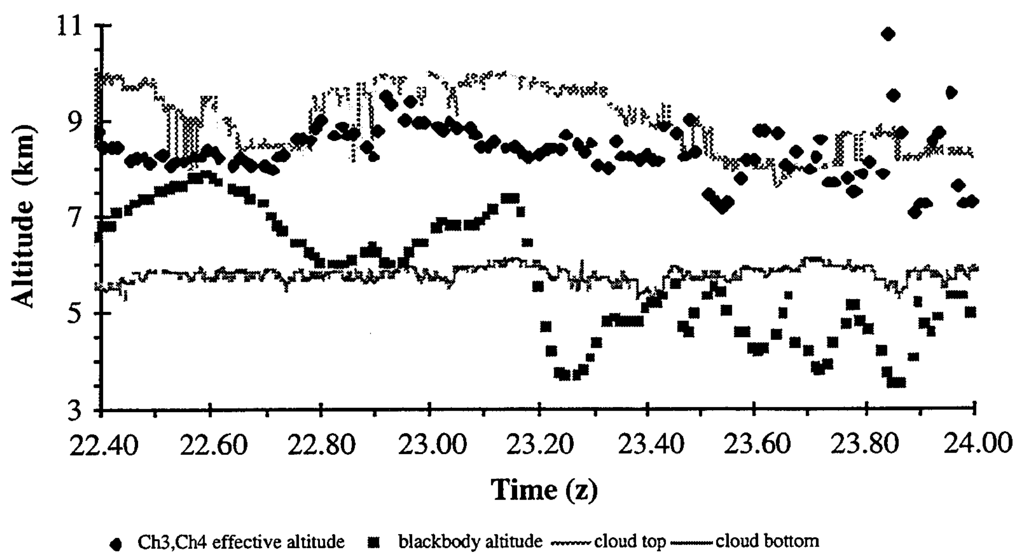


Figure 9. Enhanced Cirrus Retrievals over Hanscom AFB

The emissivity and effective altitude estimates obtained for the Hanscom cirrus sample using the cirrus analysis method are plotted in Figure 9. Pixel-by-pixel measurements of brightness temperature pairs T_3 and T_4 were selected ± 30 minutes upwind and downwind of the radar site within the cirrus clouds. The model-generated cloud altitudes for these pixels are consistent with TPQ-11 ground-based radar measurements of cirrus cloud base and top. The TPQ-11 radar measured cirrus base and tops in the 5.5 - 10 km range for the time period several hours before and after the 2337 UTC satellite overpass. The retrieved altitudes are mainly in the 8 - 10 km range, and verify well with observation.

Altitude retrievals that do not take into account the transmissive nature of cirrus, so-called "blackbody" cirrus altitudes, were also generated and compared to the satellite-derived retrievals. Blackbody altitudes are obtained by simply 1) comparing the Channel 4 thin cirrus brightness temperature directly to the atmospheric temperature profile, and 2) choosing that height whose temperature matches the observed brightness temperature. In Figure 9 are plotted the blackbody altitudes along with their corresponding satellite-derived altitudes. Discrepancies are small for the thicker cirrus clouds (where effective emissivities are close to one), on the order of a few hundred meters. However, differences between the two retrieved altitudes increase substantially as cirrus becomes more optically thin. For the very thinnest of clouds this difference in altitudes is on the order of 5 km, as can be seen in Figure 9. These results quantitatively illustrate the importance of accounting for transmissive cirrus effects when trying to retrieve accurate cirrus altitudes from satellite.

Although the size of this sample set is small, the comparison results demonstrate that it is feasible to compute accurately multichannel cirrus emissivities, transmissivities, optical depths, and effective altitudes on a pixel-by-pixel basis using the principles and techniques outlined in this report.

5. AIMS DATABASE UPGRADE - GEQS

Several major capital improvements were made to AIMS during calendar year 1995. A major aspect of these improvements was directed at upgrading the AIMS Satellite Database (ASDB). Specifically, these improvements were:

- DEC-Alpha Server 2000; 256 Mbytes RAM, 8 Gbytes of on-line disk storage
- Open-VMS operating system, C/C++/Fortran Compilers
- ORACLE Relational Database Management System (RDBMS) software

Under the scope of this project, a task was initiated to begin an upgrade of ASDB. The new system described here is called the Global Environmental Query System (GEQS). The tasks addressed include:

- Survey ASDB users and determine needs and desires for improvements
- Analyze Master Environmental Library (MEL) system concept, requirements and standards; determine applicability to GEQS
- Develop GEQS requirements
- Develop GEQS architecture
- Initiate GEQS database design

Appendix B contains a presentation covering the capabilities of the current ASDB. Upon reviewing the capabilities and limitations of ASDB, the areas requiring improvement were identified as shown in Table 12. Also indicated in this table is the extent to which ORACLE (or in general any modern relational database) supports the desired feature. Note that six of the ten features are directly supported by the ORACLE RDBMS, two are facilitated and only two others are not supported to a significant extent. Table 13 summarizes the design requirements for GEQS with those slated for incorporation in Build 1, identified with a check mark.

Table 12. Required Improvements to Current AIMS Database

REQUIRED IMPROVEMENT	SUPPORT BY ORACLE RDBMS		
	Direct	Facilitated	Other
Improved robust client-server solution for distributed network access	√		
Flexible, GUI-based general purpose interface for database queries, modifications and deletions	√		
Extensible	√		
RDBMS manages satellite data in addition to metadata	√		
Support data quality-control			√
Support query by geography			√
Minimize/eliminate maintenance of include files		√	
Minimize/eliminate kludging of database resources		√	
Ability to dynamically modify record structures	√		
Ability to pose queries across dictionary boundaries	√		
Direct = ORACLE features inherently support the desired feature; Facilitated = ORACLE provides system which simplifies including of features compared to current system; Other = features not directly related to ORACLE capabilities			

Table 13. GEQS Requirements

REQUIREMENT	BUILD 1	COMMENTS
Automatic Ingest Of AIMS Real-Time Data		
GOES-7		
GOES-8 (or -9) Imager	√	Only one at a time, now
TIROS (HRPT)		
DMSP (RTD)		
METEOSAT		
Manual Ingest Of AIMS And Other Satellite Data Sources		
All above sources		
GMS	√	
GOES-8 (or -9) Sounder	√	
TIROS (LAC/GAC)	√	
DMSP (SDS, SDF, RDS)	√	
Other	√	
Manual Ingest Of Misc. Data Sources		
Surface observations	√	
Upper air observations	√	
Other	√	
Query Types		
Date/time	√	
Sensor type	√	
Channels	√	
Processed data type, level of processing	√	
Coverage by Lat/Lon		
Platform	√	
Satellite characteristics		
Ground station status		
Data quality	√	
Data group or experiment	√	
Automatic Processing Of Data — SERCAA		
GOES-Old		
Level 1	√	Ingest/calibration, geolocation
Level 2	√	Cloud cover
Level 3		Layer and type
GOES-Next		
Level 1	√	
Level 2	√	
Level 3		
Meteosat		
Level 1-3		
TIROS		
Level 1-3		
DMSP		
Level 1-3		
SERCAA Level 4 – Merge		Merging of analyses
SERCAA Phase 2 Cirrus Properties		
GOES-Next		
TIROS		
Browse Images		Compressed images
Automatic generation		
Query and display		
API Toolkit		Interface to FORTRAN, C, C++ programs
Version 1.0	√	
Data Query GUI-Driver		Runs on ORACLE client
Version 1.0	√	Basic capability
Storage Management		
Manual storage management	√	
Storage Management GUI		
Semi-automated storage management		
Hierarchical Storage Management		
WAIS Index Generation		
Manual	√	
Automated		
Data Formats Supported		
TIFF	√	
Raster image file	√	
HDF 4.0		Proposed as standard
NITF		Incorporates various formats

5.1 GEQS SYSTEM ARCHITECTURE

Figure 10 illustrates the GEQS system architecture. The ORACLE database engine runs on the DEC Alpha Server 2000. SQL is the primary interface to the database. The SQL interface serves as the means for data query and retrieval and the Application Programmer Interface (API). These are the principal means for users to query and access data. A separate interface for the Wide Area Information Server (WAIS) is also shown. This interface is used by the Master Environmental Library (MEL) Program as described in Section 5.3.

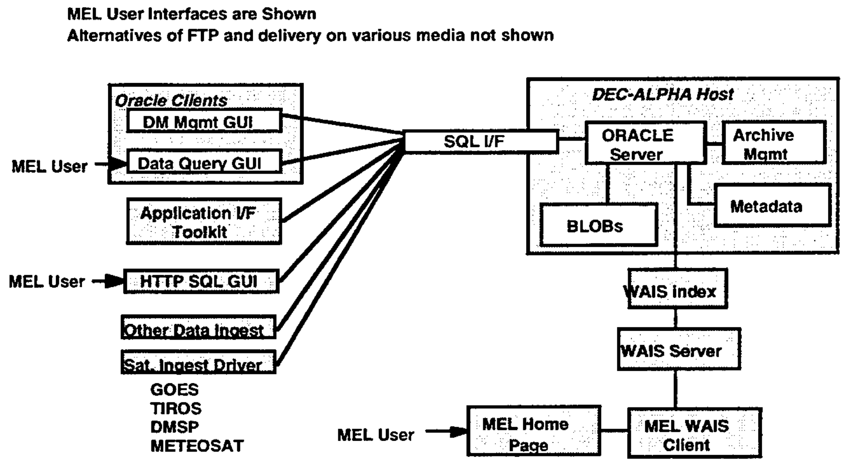


Figure 10. GEQS Architecture

One of the more critical aspects of the design implementation is the realtime, automated ingest of satellite data. This ingest must handle not only the ingest and storage of the raw satellite data, but also the calibration and higher level processing. The ingest process will require an interface between the satellite receiving workstations and the ORACLE database. Figure 11 illustrates the desired GEQS hardware/software system architecture allocations.

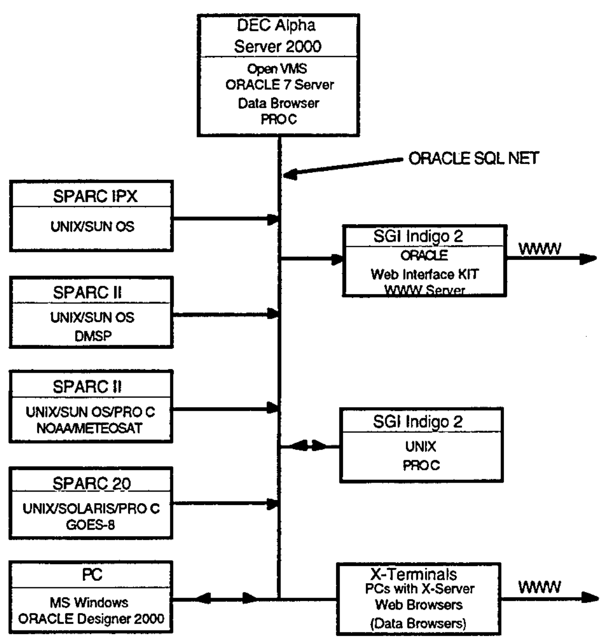


Figure 11. GEQS Hardware/Software System Architecture

5.2 RDBMS DESIGN METHODOLOGY

We are applying the Entity Relationship (ER) Method (Chen, 1983; Atzeni et al., 1993) to the design of GEQS RDBMS. While a detailed description of the ER Method is beyond the scope of this report, a brief summary follows. An entity is an object, that is distinct and distinguishable in one or more characteristics from all other objects. An entity set is a class of objects from a single category. For example: the 18Z, 9-APR-96 NOAA 12 pass is an entity from the set of NOAA satellite passes. Attributes are functions mapping an entity set into a domain (i.e., a set of allowable values). Each entity set may have zero or more attributes. For each attribute applicable to an entity set, each member of that set is mapped to one and only one domain value. Entities can have relationships among themselves. These relationships can be of the form "one-to-one", "many-to-one", "many-to-many" and "one-to-many". These relationships can be between entities in the same entity set and members of other entity sets. The basic ER Modeling process is given below:

1) Identify entities and entity sets
2) Identify attributes and domains
3) Identify relationships between entities
4) Analyze entity relationship structure and put in normal form
5) Optimize performance (storage, search, speed, other) - iterate as necessary over Steps 1-4

The goal of the penultimate step is usually to put the model in "Normal Form", although other factors may indicate deviations from a strictly normal form that are appropriate for a given application. This optimization step is iterative and almost always requires a thorough understanding of both the data and user operation of the database. For example, speed optimization may be concerned with average case or worst case performance and a knowledge of user query types is important. As an example of the "other" optimization criteria, one might choose the database so that it is easy to extend it in certain anticipated ways - such that the extensions minimize changes in system aspects.

We have selected Designer 2000 (a product of ORACLE Corporation) as the tool to support the database design and implementation. Designer 2000 will not only support the ER Modeling Process, but it will also produce the SQL code to directly configure the ORACLE RDBMS.

One key design decision for GEQS is the storage method for the satellite imagery itself. The current AIMS database stores this information separately from the database with the main database simply identifying the files names and storage locations of the particular image data. Modern databases support what are known as Binary Large OBjects or BLOBs. BLOBs are handled as database fields and can be accessed via the standard SQL interface. This offers considerable simplification compared to a system which must arrange a separate set of software calls to handle an actual satellite data file once it is identified. The full power of the RDBMS is also available, particularly with regard to restructuring, etc. A possible disadvantage of the BLOB approach to the storage of satellite data is efficiency.

We intend to prototype our initial database based on storage of the imagery data as BLOBs. We will then investigate the efficiency, power and flexibility tradeoffs before making a final determination as to how the final database should be configured.

5.3 MASTER ENVIRONMENTAL LIBRARY

One of the goals of GEQS is to integrate with the Master Environmental Library (MEL). MEL is intended to be used as a central facility from which members of the simulation, modeling, and research communities can survey available data for data likely to support their need. MEL does not itself host the data. Rather, regional and other allied sites are the repositories of this data. Once potentially applicable data are identified through the initial MEL search mechanism (see below), users then interface directly with the appropriate sites. A more detailed survey of the data could take place at the site and actual data delivery would then be conducted with mechanisms specified by that site.

MEL has specified that the initial MEL-directed search be conducted using the Wide Area Information Server (WAIS). WAIS uses client-server protocols based on a server accessing a specially indexed data set. WAIS clients then can search this index for one or more key words and with various boolean-type search options.

We plan to work with the MEL Project staff to develop an AIMS/GEQS WAIS-searchable index. This would be the first step towards making AIMS a regional MEL site. As the capabilities of AIMS/GEQS expand and MEL continues to develop, we would continue to work with MEL to further integrate AIMS into their system.

5.4 GEQS STATUS AND PLANS

We have completed the initial requirements review and developed the baseline GEQS architecture. The goal of this effort is to develop a system that fulfills the capabilities designated in Table 13 for Build 1. With current authorized funding, we plan to perform the following tasks:

- Complete a Software Requirements Document
- Finalize documentation of design approach/concept
- Complete initial database design using Designer 2000

Contingent on continued funding for this task, the following tasks could be performed in the next year:

- Develop and populate initial database
- Develop Application Programmer Interface (API) toolkit
- Develop and implement automated satellite data ingest for GOES
- Develop and implement interface for automated SERCAA GOES Level 1 and 2 processing
- Develop initial version of database search GUI
- Produce WAIS index and interface to MEL WAIS server

6. REFERENCES

Atzeni, P. and V. De Antonellis, 1993: Relational Database Theory. Benjamin Cummings Publishing Co., 389 pp.

Barton, I. J., 1983: Upper Level Cloud Climatology from an Orbiting Satellite. Journ. Atmos. Sci., 40, 435 - 447.

Chen, P.P. (Editor), 1983: Entity-Relationship Approach to Information Modeling and Analysis. North-Holland, Amsterdam, the Netherlands.

d'Entremont, R. P., M. K. Griffin, and J. T. Bunting, 1990: Retrieval of Cirrus Radiative Properties and Altitudes Using Multichannel Infrared Data. Proceedings, Fifth Conference on Satellite Meteorology and Oceanography, Amer. Meteor. Soc., 3-7 September 1990, London, England, 4-9.

d'Entremont, R. P., D. P. Wylie, S.-C. Ou, and K.-N. Liou, 1996: Retrieval of Cirrus Radiative and Spatial Properties Using Coincident Multispectral Imager and Sounder Data. Preprints, Eighth Conf. on Satellite Meteorology and Oceanography, Amer. Meteor. Soc., Atlanta, GA, 28 January - 2 February 1996, 377-381.

Gustafson, G.B., R.G. Isaacs, R.P. d'Entremont, J.M. Sparrow, T.M. Hamill, C. Grassotti, D.W. Johnson, C.P. Sarkisian, D.C. Peduzzi, B.T. Pearson, V.D. Jakabhazy, J.S. Belfiore, and A.S. Lisa, 1994: Support of Environmental Requirements for Cloud Analysis and Archive (SERCAA): Algorithm Descriptions. PL-TR-94-2114, Phillips Laboratory, Hanscom AFB, MA, ADA283240.

Hoke, J.E., J.L. Hayes, L.G. Renninger, 1981: Map projections and grid systems for meteorological applications. AFGWC-TN-79-003, Air Weather Service, Scott AFB, IL.

Hunt, G.E., 1973: Radiative Properties of Terrestrial Clouds at Visible and Infra-Red Thermal Window Wavelengths. Quart. Journ. Royal Meteor. Soc., 99, 346 - 369.

Menzel, W. P., D. P. Wylie, and K. I. Strabala, 1992: Seasonal and Diurnal Changes in Cirrus Clouds as Seen in Four Years of Observations with the VAS. Journ. Appl. Meteor., 31, 370 - 385.

Ou, S. C., K. N. Liou, W. M. Gooch, and Y. Takano, 1993: Remote Sensing of Cirrus Cloud Parameters Using AVHRR 3.7 and 10.9 mm Channels. Appl. Optics, April 1993 (in press).

Parol, F., J. C. Bruiz, G. Brogniez, and Y. Fouquart, 1991: Information Content of AVHRR Channels 4 and 5 with respect to the Effective Radius of Cirrus Cloud Particles. Journ. Appl. Meteor., 30, 973 - 984.

Schiffer, R. A., and W. B. Rossow, 1983: The International Satellite Cloud Climatology Project(ISCCP): the First Project of the World Climate Research Programme. Bull. Amer. Meteor. Soc., 64, 779 - 784.

Smith, W. L., H. E. Revercomb, H. B. Howell, and M.-X. Lin, 1990: Multi-spectral Window Radiance Observations of Cirrus From Satellite and Aircraft, November 2, 1986 Project FIRE. FIRE Science Results 1988, NASA Langley Res. Ctr., 89-93.

Takano, Y., and K. N. Liou, 1989: Solar Radiative Transfer in Cirrus Clouds. Part I: Single Scattering and Optical Properties of Hexagonal Ice Crystals. Journ. Atmos. Sci., 46, 3 - 19.

Woodbury, G. E., and M. P. McCormick, 1983: Global Distribution of Cirrus Clouds Determined from SAGE Data. Geophys. Res. Lett., 10, 1180 - 1183.

Wylie, Donald P., and W. P. Menzel, 1989: Two Years of Cloud Cover Statistics Using VAS. Journ. of Clim., 2, 380 - 392.

APPENDIX A

DNA DATA SAVE DOCUMENTATION REPORTS

P601

Data Save Documentation Report No. 1

ADVANCED GEOPHYSICAL ENVIRONMENT SIMULATION TECHNIQUES

Task 1: Satellite Data Sets for Worldwide Cloud Prediction

This data documentation report covers initial data set generation
for three SERCAA regions of interest:

Japan, Himalayas, and Panama

for the period:

27 - 30 May 1993

Contract Number F19628-94-C-0106

issued by:

Electronic Systems Division
Air Force Systems Command
Hanscom AFB, MA 01731

Submitted by:

Atmospheric and Environmental Research, Inc.
840 Memorial Drive
Cambridge, MA 02139

14 October 1994

David B. Hogan
Gary B. Gustafson
Principal Investigators

1.0 Introduction

This Data Documentation Report provides a description of the first data save made in accordance with the revised statement of work for Satellite Data Sets for Worldwide Cloud Prediction Models. It is intended to provide a description of the data, format, how it was gathered and processed, and a description of the algorithms used to generate it. The data set consists of raw satellite data and analyzed products produced by the SERCAA cloud analysis algorithms. The period covered is 27-30 May 1993 for three SERCAA regions of interest: Japan, Himalayas (Bangladesh), and Panama. The regions of interest have the following (i,j) 16th mesh grid coordinates: Japan (344,234 - 408,298), Himalayas (536,168 - 600,232), Panama (504,914 - 568,978). All available data from those dates are included. These data were processed as a part of the SERCAA validation program in the spring of 1994, and not specifically for DNA. As such data formats may not be in the final form that subsequent DNA data sets will be provided in although every effort has been made to anticipate future requirements and include them in the current data formats.

2.0 Processing Environment

SERCAA satellite data processing for this data set used the cloud analysis algorithms described by Gustafson et al. (1994). Multisource data from the DMSP F10 and F11, NOAA-11 and NOAA-12, GMS-4, and GOES-7 satellites were used. Data sources were as follows: DMSP - National Geophysical Data Center (NGDC), Boulder, CO; NOAA - National Climatic Data Center (NCDC), Ashville, NC; GMS - Sea Space Corp., San Diego, CA; GOES - Phillips Laboratory. All data except GOES were received on tape in various formats. GOES data were received live at the AIMS ground station. Data processing was performed on the Air Force Interactive Meteorological System (AIMS) and at the AER computer center in Cambridge, MA. Four levels of data processing are performed by the SERCAA cloud analysis algorithms as summarized in Figure 1.

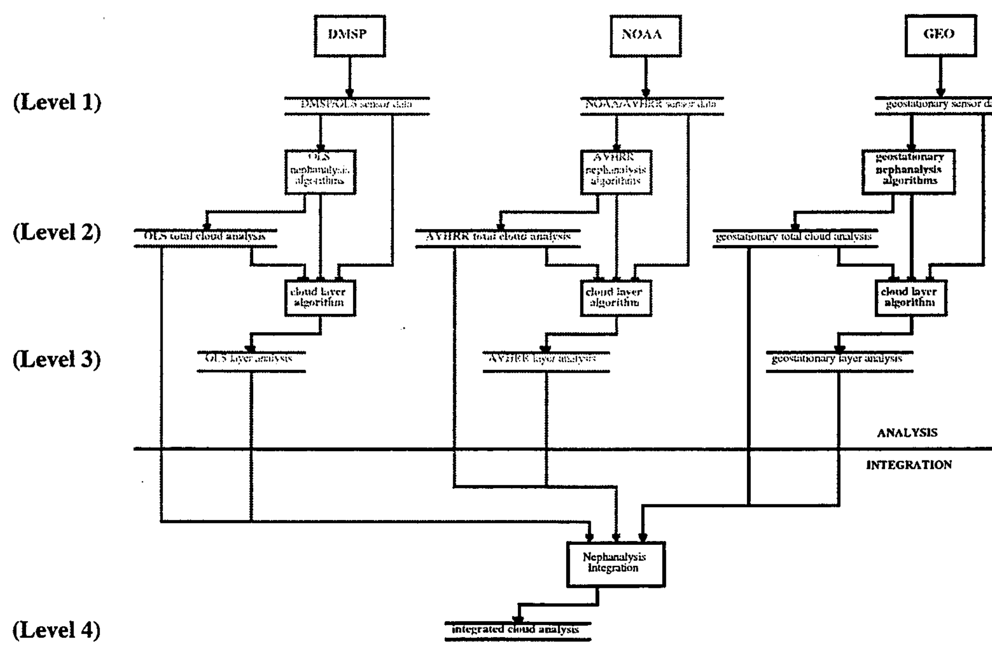


Figure 1 SERCAA data flow and processing levels

Level 1 processing consists of data ingest. Tape and live data are processed through separate ingest programs depending on the data source and format. All data are then stored in a standard format in the original satellite scan projection (i.e., flat files where the number of elements correspond to the number of pixels in a scan line and the number of rows corresponds to the number of scan lines) that are maintained on AIMS through the SERCAA Database (SDB) management software. Level 1 data products consist of separate files for each sensor channel plus files containing Earth location information and satellite/solar geometry information. Satellite data characteristics are summarized in Table 1. In cases where visible and infrared channel resolution differs, the higher resolution data are subsampled to match the coarser resolution data (e.g., GMS visible data are subsampled by a factor of four to match the IR data resolution). Earth location data consist of latitude-longitude pairs that are maintained at a subsampled resolution relative to the satellite data. For each sensor scan line, one latitude-longitude pair is provided for every n^{th} pixel, where n varies with satellite. Geometry information are also subsampled in exactly the same ratio as the Earth location information and consist of three angles: satellite zenith, solar zenith, and sun-satellite azimuth. Ingest products are described more completely in Section 2 of Gustafson et al. (1994).

Table 1. Sensor Channel Data Attributes During SERCAA

Satellite	Sensor	Channel (µm)	Data Format	Resolution[1] (km)	Bits per Pixel[2]	Pixels per Scan Line
DMSP	OLS	0.40-1.10	counts	2.7	6	1464
		10.5-12.6	EBBT	2.7	8	1464
NOAA	AVHRR	0.58-0.68	percent albedo	4.0	10	409
		0.72-1.10	percent albedo	4.0	10	409
		3.55-3.93	EBBT	4.0	10	409
		10.3-11.3	EBBT	4.0	10	409
		11.5-12.5	EBBT	4.0	10	409
GOES	VAS	0.55-0.75	counts	0.86	6	15288
		3.71-4.18	EBBT	13.8	10	1911
		10.5-12.6	EBBT	6.9	10	3822[3]
		12.5-12.8	EBBT	13.8	10	1911
METEOSAT	VISSR	0.55-0.75	counts	2.5	8	5000
		10.5-12.6	EBBT	5.0	8	2500
GMS	VISSR	0.5-0.75	counts	1.25	6	10000
		10.5-12.5	EBBT	5.0	8	2500

[1]Sensor resolution at satellite subpoint that will provide global coverage.

[2]AVHRR radiance data are transmitted at 10-bit resolution, however, the SERCAA development system could only accommodate 8-bit brightness temperature data (although the full 10-bit resolution is used in the radiance to brightness temperature transformation).

[3]GOES long wave infrared data are over sampled in the across-track direction by a factor of 2.

Level 2 processing consists of sensor specific nephanalysis algorithms. Level 1 sensor data from DMSP, NOAA, and the geostationary satellites are processed through separate nephanalysis algorithms as indicated in Figure 1. Each time data from a new satellite pass are ingested, the appropriate nephanalysis algorithm is run and results are placed in a Level 2 output file. One output file is generated for each nephanalysis run and nephanalysis results are stored in the original satellite scan projection with one byte of information for each pixel. Each byte is bit packed according to the map in Table 2.

Table 2. Cloud Analysis Algorithm MCF File Bit Assignments

Bit	Assignment	Description
0	Cloud Mask	ON = Cloud-Filled OFF = Cloud-Free
1	Low Cloud	ON = Low Cloud
2	Thin Cirrus Cloud	ON = Thin Cirrus Cloud
3	Precipitating Cloud	ON = Precipitating Cloud
4	Partial Cloud	Only used by geostationary algorithm
5	Data Dropout	ON = Missing or Unreliable Data
6	Confidence	0 = Missing Data; 1 = Low;
7	Flag	2 = Middle; 3 = High

Level 3 processing uses Level 1 and 2 products as input to segment the scene into vertical cloud layers and to classify different cloud types. It also remaps the data from the individual satellite projections to the AFGWC standard polar stenographic map projection at 16^{th} mesh grid resolution (Hoke et al., 1981). The three regions of interest processed for the May 1993 data set have the following (i,j) 16^{th} mesh grid coordinates: PAN (504,914 - 568,978), HIM (536,168 - 600,232), JPN (344,234 - 408,298). For each grid cell the information in Table 3 is provided. Note: no Level 3 products are available for the May 1993 data save.

Table 3. Cloud Typing and Layering Output

Column	Parameter	Description
1		i-Coordinate for Grid Cell
2		j-Coordinate for Grid Cell
3		Layer of Grid Cell for Which the Statistics Pertain
4	(CTT)	Cloud Top Mean IR Temperature of Pixels in Layer
5	(CTTV)	Cloud Top IR Temperature Variance of Pixels in Layer
6	(NPL)	Total Number of Pixels in Layer
7	(NPIX)	Total Number of Pixels in Grid Cell
8	(IDD)	Total Number of Data Dropouts in Grid Cell
9	(TYP)	Cloud Type of Layer
10	(ICF)	Mean Confidence Flag for Layer
11	(LCC)	Total Number of Low Cloud Pixels Detected in Cloud Analysis
12	(TCC)	Total Number of Thin Cirrus Pixels Detected in Cloud Analysis
13	(PCC)	Total Number of Precipitating Cloud Pixels Detected in Cloud Analysis
14	(PTC)	Total Number of Partial Cloud Pixels Detected in Cloud Analysis

Level 4 processing operates on the most recent Level 3 products available from each satellite source. One new Level 4 integrated analysis is performed each hour. Thus while Level 1, 2, and 3 products are event driven (i.e., resulting from the ingest of a new satellite pass), Level 4 processing is schedule driven (i.e., one analysis/hour). The Level 4 output parameters are summarized in Table 4. The output format differs from the previous levels in that gridded fields of each parameter are output separately as opposed to combining all parameters for each grid cell into one structure. For example, for the first output parameter in Table 4, Number of Cloud Layers (NLAY), an array is

dimensioned to the full size of the output grid (NX*NY), populated with the NLAY values for each grid cell, and then output as one record in the output file. Note: for the May 1993 data, all output grids are the same size, 65X65 grid cells. The next parameter (CFT) would be the next output record, and so on.

Table 4 Analysis Integration Processed Parameters

Parameter	Description	Dimensions	Integrated Product
NLAY	Number of Cloud Layers	NX*NY	Yes
CFT	Total Cloud Fraction	NX*NY	Yes
CF	Layer Cloud Fraction	NX*NY*NZ	Yes
CTT	Layer Cloud Top IR Temperature	NX*NY*NZ	Yes
ICF	Analysis Confidence Flag Index	NX*NY*NZ	Yes
ITY	Layer Cloud Type	NX*NY*NZ	Yes
ECFT	Estimated Error in Total Cloud Fraction	NX*NY	Yes
ECF	Estimated Error in Layer Cloud Fraction	NX*NY*NZ	Yes
CTTSD	Local Standard Deviation of Analyzed Cloud Top IR Temperature	NX*NY*NZ	No
ICB	Precipitating Cloud Detection Index	NX*NY*NZ	No
ICI	Thin Cirrus Cloud Detection Index	NX*NY*NZ	No
ILO	Low Cloud Detection Index	NX*NY*NZ	No

NX=number of columns in analysis grid (65)
NY=number of rows in analysis grid (65)
NZ=maximum number of layers (4)

3.0 Tape Format

All data for the May 1993 data save are contained on two 8 mm tapes written in UNIX tar format. The first tape, labeled: SERCAA RE May 93 1, contains all the Level 1 and 2 products. The second tape, labeled: SERCAA IA May 93 1, contains all Level 4 products. The size of the combined Level 1 and 2 products is approximately 601 Mbytes and the Level 4 products occupy 26 Mbytes. In addition to the two tapes, hard copy listings of the contents for both tapes are also provided. The hard copy listings are required to locate specific data sets on the tapes.

Level 1, 2, and 3 products are generated for each new pass of satellite data received for the period of the data save. For data archiving purposes all Level 1 and 2 products associated with a given satellite pass are placed in a single directory and subsequently placed on tape as a single tar file. Thus the first tape contains a series of tar files that each correspond to all Level 1 and 2 products associated with a single satellite pass. Level 4 files are grouped on tape by day, thus for the May data save there are four tar files on the Level 4 tape that each contain all Level 4 output files for each of the four days 93147-93150 (27-30 May 1993). For each set of Level 1 and 2 products, and for each Level 4 file there is also an SDB Information File. These files contain descriptive metadata information extracted from the SERCAA Database that describe the relevant attributes of the SERCAA product files. For example, information files list the number of pixels in a scan line of satellite data and the number of scan lines in the file.

Information on subsampling ratios for the Earth location and angles files are also contained there.

Detailed descriptions of the file formats used for each output level, and the associated information files, provided for the May 1993 save (Level 1, 2, and 4) are described in detail in Appendix A. Level 3 file formats are still under development and will be provided at a later date but prior to delivery of the March 1993 data save due on 23 November 1994. Appendix B provides a guide for extracting data sets from tape.

4.0 References

Gustafson, G.B., R.G. Isaacs, R.P. d'Entremont, J.M. Sparrow, T.M. Hamill, C. Grassotti, D.W. Johnson, C.P. Sarkisian, D.C. Peduzzi, B.T. Pearson, V.D. Jakabhazy, J.S. Belfiore, and A.S. Lisa, 1994: Support of Environmental Requirements for Cloud Analysis and Archive (SERCAA): algorithm descriptions. PL-TR-94-2114, Phillips Laboratory, Hanscom AFB, MA, ADA283240.

Hoke, J.E., J.L. Hayes, L.G. Renninger, 1981: Map projections and grid systems for meteorological applications. AFGWC-TN-79-003, Air Weather Service, Scott, AFB, IL.

Appendix A

Archive Data Format Descriptions

By Level

Level 1: Satellite Image Files

Satellite image filenames as they appear on tape have the following naming convention:

SSS_CCC_ROI_DDD_HH.Dat or .Tif

where

SSS - Satellite identifier:

F10	DMSP F-10
F11	DMSP F-11
N11	NOAA-11
N12	NOAA-12
G04	GMS (Japan)
G07	GOES-7 (U.S.)

CCC - spectral channel identifier

ROI - Region of Interest:

SAT for the North America test bed
SET for the Asia test bed
SDT for the Africa test bed

DDD - Julian day for which the image data are valid

HH - UTC hour of image data

Dat - Raw image file format

Tif - TIFF file format

The exception to this naming convention is the naming of GOES-7 (G07) files. G07 files use the following naming convention:

CCCHHMM.Dat or .Tif

where

CCC - Spectral channel identifier:

VIS for visible data
IR1 for infrared channel 1
IR2 for infrared channel 2
IR3 for infrared channel 3

HH - UTC hour of image data

MM - UTC minute of image data

File and Record Structure

All image files contain fixed-length records. The number of lines and number of elements in an image file ARE contained in the Related Entries (RE) SDB information file, under the "[SATIMG]" heading:

NUM_LINES	Number of image data lines in the file.
ELEM_PER_LINE	Number of elements (pixels) per line.
BYTES_PER_ELEMENT	Number of bytes per pixel. This number is 1 for all SERCAA imager sensor data.

Image file data are stored in Tagged Image File Format (TIFF), therefore an alternative way to determine image dimensions is to read the TIFF header and examine the width and height fields.

Image pixel values represent either counts or albedo for visible data, and brightness temperatures for thermal infrared data. The following table summarizes the attributes of the SERCAA image data values.

Table 1

Satellite ID (SSS)	Spectral Channel (CCC)	Channel Type	Wavelength Band	Physical Value
F10 or F11	001	Visible	0.4 - 1.1 μm	Counts[1]
	002	Long-Wave IR	10 -12 μm	Brightness Temp.[2]
N11 or N12	001	Visible	0.63 μm	Albedo[3]
	002	Near-IR	0.86 μm	Albedo
	003	Mid-Wave IR	3.7 μm	Brightness Temp.
	004	Long-Wave IR	10.7 μm	Brightness Temp.
	005	Long-Wave IR	11.8 μm	Brightness Temp.
G04	001	Visible	0.55 - 0.75 μm	Counts
	002	Long-Wave IR	10.2 - 11.2 μm	Brightness Temp.
G07	VIS	Visible	0.55 - 0.75 μm	Counts
	IR1	Long-Wave IR	10.5 - 12.6 μm	Brightness Temp.
	IR2[4]	Long-Wave IR	13.33 μm	Brightness Temp.
		Long-Wave IR	12.5 - 12.8 μm	Brightness Temp.
		Infrared	6.7 μm	Brightness Temp.
		Infrared	7.25 μm	Brightness Temp.
	IR3[5]	Infrared	6.7 μm	Brightness Temp.
		Mid-Wave IR	3.71 - 4.18 μm	Brightness Temp

[1]Visible counts range from 0 - 255. High counts denote highly reflective surfaces and low counts denote poorly reflective surfaces.

[2]Brightness temperatures are byte-encoded such that the range 0 - 255 corresponds to the temperature range 327.5 K to 200.0 K. The relation between byte values and temperature is linear over this range; the conversion from byte value B to brightness temperature T is given by the relation

$$T = -0.5B + 327.5.$$

[3]Albedo values are byte-encoded such that the range 0 - 255 corresponds to the albedo range 0 - 100%. The relation between byte values and percent albedo is linear; the conversion from byte value B to percent albedo A is given by the relation

$$A = 0.392B.$$

[4]GOES-7 alternates this channel with one of four bands per ingest period. The only band utilized from this channel was the 12.5 - 12.8 μm band.

[5]GOES-7 alternates this channel with one of two bands per ingest period. The only band utilized from this channel was the 3.71 - 4.18 μm band.

Level 1: **Latitude-Longitude File**

Latitude-longitude filenames as they appear on tape have the following naming convention:

SSS_LAT_ROI_DDD_HH.Dat

where

SSS - Satellite identifier:

F10	DMSP F-10
F11	DMSP F-11
N11	NOAA-11
N12	NOAA-12
G04	GMS (Japan)
G07	GOES-7 (U.S.)

LAT - a constant that identifies the file as an latitude-longitude file

ROI - Region of Interest for which the latitude-longitude file is valid:

SAT for the North America test bed
SET for the Asia test bed
SDT for the Africa test bed

DDD - Julian day of satellite data for which the Earth locations are valid

HH - UTC hour of the satellite data for which the Earth locations are valid

File and Record Structure

Latitude-longitude Earth location files contain fixed-length records, the number and size of which depend on both the size of the corresponding image files and the satellite type. There is always one latitude-longitude record corresponding to each satellite image file record, where a satellite image file record contains one image scan line of information.

The information necessary for interpreting an latitude-longitude file record is contained in the Related Entries (RE) SDB information file, under the "[LATLON]" heading:

LL_REC_LEN	Record length in bytes.
LL_LINE_INTERVAL	The number of image file records per lat-lon record. For the May 1993 data set this number is always 1.
LL_ELEM_INTERVAL	The subsampling rate of lat-lon information relative to the corresponding satellite data. For example, if LL_ELEM_INTERVAL = 40, there is one latitude-longitude pair for every 40th image pixel in the scan line (i.e., for pixels 1, 41, 81, ...). Linear interpolation is required to retrieve Earth location information for intermediate pixels 2-40, 42-80, ...
LL_ELEM_PER_LINE	This is the number of latitude-longitude elements per latitude-longitude file record.

A latitude-longitude file data element is a 4-byte structure that contains the scaled latitude and longitude for a given pixel. Thus the length of an latitude-longitude file record in bytes is given by:

LL_REC_LEN = 4 * LL_ELEM_PER_LINE

The 4 bytes consist of two 16-bit integer variables: LONG and LAT. The storage convention is as follows:

LONG	Pixel longitude * 128. To obtain the floating-point longitude, FLONG = LONG / 128. Longitude range is -180° to 180°, positive east.
LAT	Pixel latitude * 128. to obtain floating-point latitude, FLAT - LAT / 128. Latitude range is -90° to 90°, positive north.

Level 1: Angles File

The angles filenames as they appear on tape have the following naming convention:

SSS_ANG_ROI_DDD_HH.Dat

where

SSS - Satellite identifier:

F10	DMSP F-10
F11	DMSP F-11
N11	NOAA-11
N12	NOAA-12
G04	GMS (Japan)
G07	GOES-7 (U.S.)

ANG - a constant that identifies the file as an angles file

ROI - Region of Interest for which the angles file is valid:

SAT for the North America test bed

SET for the Asia test bed

SDT for the Africa test bed

DDD - Julian day of satellite data for which the angles are valid

HH - UTC hour of the satellite data for which the angles are valid

File and Record Structure

Angle files contain fixed-length records, the number and size of which depend on both the size of the corresponding image files and the satellite type. There is always one angles record corresponding to each satellite image file record, where a satellite image file record contains one image scan line of information.

The information necessary for interpreting an angles file record is contained in the Related Entries (RE) SDB information file, under the "[ANGLES]" heading:

ANG_REC_LEN	Record length in bytes.
ANG_LINE_INTERVAL	The number of image file records per angles record. This number is almost always 1.
ANG_ELEM_INTERVAL	The subsampling rate of angles information relative to the corresponding satellite image. For example, if ANG_ELEM_INTERVAL = 8, there is one set of angles valid for every eighth image pixel in the scan line (i.e., for pixels 1, 9, 17, 25, ...). Linear interpolation is required to retrieve angles information for intermediate pixels 2-8, 10-16, 18-24, ...
ANG_ELEM_PER_LINE	This is the number of angles elements per angles file record.

An angles file data element is a 12-byte structure containing three angles that define the satellite and solar viewing geometry for a given pixel. Thus the length of an angles file record in bytes is given by:

ANG_REC_LEN = 12 * ANG_ELEM_PER_LINE

The 12 bytes consist of three 32-bit floating-point variables: SATZEN, SOLZEN, and AZIMUTH. Note: Angle files were generated on a VMS computer. To interpret these floating-point numbers on a UNIX machine it is necessary to convert from VMS to IEEE floating-point formats. Most UNIX operating systems provide a utility to perform this conversion. Angle measurement conventions are as follows:

SATZEN	Scene satellite zenith angle, 0° - 90°.
SOLZEN	Scene solar zenith angle, 0° - 90°.
AZIMUTH	Relative angle between the solar and satellite azimuth angles, 0° - 359°. When AZIMUTH = 0°, the sun is directly behind the satellite (i.e., the viewed point, the satellite, and the sun are collinear). When AZIMUTH = 180°, the satellite is looking directly into the sun (the satellite squints to compensate).

Level 2: **Nephanalysis Products**

Nephanalysis products are stored as bit-encoded byte values known as MCF (cloud Mask and Confidence Flag). MCF filenames as they appear on tape have the following naming convention:

SSS_MCF_ROI_DDD_HH.Dat or .Tif

where

SSS - Satellite identifier:

F10 DMSP F-10
F11 DMSP F-11
N11 NOAA-11
N12 NOAA-12
G04 GMS (Japan)
G07 GOES-7 (U.S.)

MCF - a constant that identifies the file as an MCF file
ROI - Region of Interest for which the product is valid:

SAT for the North America test bed
SET for the Asia test bed
SDT for the Africa test bed

DDD - Julian day for which product is valid
HH - UTC hour for which product is valid
Dat - Raw product file format
Tif - TIFF file format

File and Record Structure

The MCF product files contain fixed-length records, the number and size of which depends on both the size of the corresponding image files and the satellite type. The following table specifies how to determine the record size and number of records in an MCF file.

Let NCOLS and NROWS be the number of columns and rows, respectively, in the corresponding satellite image file. Then:

If the image satellite id is:	Then the MCF file record size is:	And the number of records in the MCF file is:
F11 or F12	MOD(NCOLS, 16)	MOD(NROWS, 16)
N11 or N12	MOD(NCOLS, 32)	MOD(NROWS, 32)
G04 or G07	NCOLS	NROWS

The MCF file is stored in Tagged Image File Format (TIFF), therefore an alternative way to determine file dimensions is to read the TIFF header and examine the width and height fields.

The format of an MCF file is the same regardless of the satellite platform it was derived from. The first byte of the first record of the MCF file corresponds to the first byte of the first record in the corresponding image data file. As can be seen in the above table, the MCF and image file sizes are not always the same. However, the two files are always aligned with respect to the upper-left corner of each.

There is one MCF byte per analyzed image pixel. MCF bytes are bit-packed according to the following convention:

Bit 0 (least significant) is the cloud/no-cloud bit. If bit 0 is off, the corresponding image pixel is clear; if bit 0 is on, it is completely cloudy.

Bit 1 is the low cloud bit. If bit 1 is on, the pixel contains low cloud as determined by an appropriate spectral (or other) signature test.

Bit 2 is the thin cirrus cloud bit. If bit 2 is on, the pixel contains cirrus as determined by an appropriate spectral (or other) signature test.

Bit 3 is the cumulonimbus bit. If bit 3 is on, the pixel contains thunderstorm clouds.

Bit 4 is the partly cloudy bit. If bit 4 is on, the pixel is partly cloudy. If bit 4 is on, bit 0 is off. DMSP data is used exclusively to determine partly cloud conditions.

Bit 5 is the bad data bit. It is set whenever satellite data are missing or unreliable. If set, all other bits should be ignored.

Bits 6 and 7 contain the confidence level attached to the accuracy of the cloud/no-cloud decision for the corresponding cloudy image pixel. Confidence levels are rated as 0 for missing data, 1 for low confidence, 2 for mid-level confidence, and 3 for high confidence.

Low cloud, thin cirrus, and cumulonimbus conditions are always associated with completely cloud conditions (i.e., bit 0 will always be on in the presence of one or more of these conditions). Cloud level and cloud type are not detected under partly cloudy conditions (i.e., if bit 4 is on, bits 1 through 3 will be off).

Example:

MCF byte	1 1 0 0 0 1 0 1	(C5 in hex)
bit position	7 6 5 4 3 2 1 0	

The corresponding image pixel contains thin cirrus that has been detected with a high level of confidence.

Level 4: **Integrated Product**

The integrated product filenames as they appear on tape have the following naming convention:

ALL_IAN_ROI_DDD_HH.Dat

where

ALL and IAN are constants (Integrated ANalysis from ALL sensors)
ROI - Region of Interest for which the product is valid:
PAN for the Panama ROI
BAN for the Bangladesh ROI
JAP for the Japan ROI
DDD - Julian day for which product is valid
HH - GMT hour for which product is valid

File Structure

For the May 1993 data save, all integrated product files contain 24 fixed-length 4225 byte records. Table 1 summarizes the contents of each record. Except for the header record (described below), each record contains data values for one of the parameters listed in Table 1.

Table 1. Integrated Product File Structure

Rec. No.	Field	Units	Range	Missing or bad value	Scaling	Comments
1						Header record See Table 3
2	Total cloud fraction	Percent	0 - 100	255		
3-6	Cloud fraction by layer	Percent	0 - 100	255		
7	Number of cloud layers		0 - 4	255		
8-11	Cloud top temperature by layer	k	0 - 127	255	T - 200	Unscaled Range: 200-327 K
12-15	Cloud type by layer		0 - 9			See Table 2
16	Total cloud error	Percent	0 - 100	255		
17-20	Total cloud error by layer	Percent	0 - 100	255		
21-24	Confidence flags by layer		10 - 30			Low to high confidence

Table 2. Cloud Type Codes

Cloud Type Code	Cloud Type
0	No Cloud
1	Cirrus
2	Cirrostratus
3	Altocumulus
4	Altostratus
5	Stratocumulus
6	Stratus
7	Cumulus
8	Cumulonimbus
9	Nimbostratus

Record Structure

Each record contains data values valid for grid points within a 65 X 65 2-D grid. The grid is superimposed on a hemispheric secant polar stereographic map projection. Grid resolution is based on a whole mesh grid spacing of 381 km at 60^{o} latitude and nested grids are defined in terms of the number of grid cells that fit within a whole mesh grid. The integrated product grid is a 16^{th} mesh grid (i.e., 16 X 16 cells per whole mesh box). The collection of data values (for a particular parameter) over the entire 2-D grid is referred to in this documentation as a field. Note that 65 X 65 X 1 byte/datum = 4225 bytes/record.

Header Record Structure

All values are 16-bit integers. Refer to Table 3.

Table 3. Header Record Structure

Field	Description	Range	Comments
1	Analysis type		always = 0
2	Year of product in YY format		
3	Julian day of product in DDD format		
4	GMT time of product in HHMM format		
5	First guess Boolean	0 or 1	0 = no first guess used 1 = first guess used
6	Year of first guess		
7	Julian day of first guess		
8	GMT time of first guess		
9	No. of satellite analyses used to derive product	1 to 3	
10	Flag for GOES data usage	0 or 1	A non-zero value indicates data was used.
11	Flag for NOAA data usage	0 or 2	" "
12	Flag for DMSP data usage	0 or 3	" "
13	Flag for GMS data usage	0 or 4	" "
14	Flag for M5 data usage	0 or 5	" "
15	Year of GOES data		If = 999 data were not used. Same for fields 16-29.
16	Julian day of GOES data		
17	GMT time of GOES data		
18	Year of NOAA data		
19	Julian day of NOAA data		
20	GMT time of NOAA data		
21	Year of DMSP data		
22	Julian day of DMSP data		
23	GMT time of DMSP data		
24	Year of GMS data		
25	Julian day of GMS data		
26	GMT time of GMS data		
27	Year of Meteosat data		
28	Julian day of Meteosat data		
29	GMT time of Meteosat data		
30	Starting 16th mesh column no. of ROI		
31	Starting 16th mesh row no. of ROI		

The header record is padded with 4163 bytes to complete a 4225 byte record.

Appendix B

Data Extraction Guide

****SERCAA DATA SET RELEASE TO DNA****

****
What should I have ?

DNA_RELEASE.TXT	This document.
(2) 8 mm D8-112 tapes	One tape contains the SERCAA Integrated Analysis (SIA) data files. (Aprox 26 MB). The other tape contains the Related Entry (RE) data (which consists of Satellite, Latitude/Longitude, Angles(Geometry) and Product(cloud mask) data files. (Aprox 601 MB).
SIA.TAR.LIST	A hard copy of the file that lists the tar contents of the SIA data files.
ENTRIES.TAR.LIST	A hard copy of the file that lists the tar contents of the RE data files.
DATA_DESCRIPTION	A hard copy document that describes the specifics of the various data types.

****
What type of tape drive was used ?

A SUN Exabyte EXB-8500 8 mm tape drive recording in high density mode (5 gig).

****
What utility was used to create the release tapes ?

The data were placed on the tapes using a SUN SPARC II running SUN OS 4.1.2. The following tar command syntax was used:

```
sun% tar cvBf /dev/nrst8 somedirectory
```

****
How are the data arranged on the release tape ?

The data on the SIA tape are contained in four tar files. Each of these tar files represents a directory that contains all the SIA data for a particular day (day 93147 through day 93150). Each directory name follows the convention:

CYYJJJ

where:

C = century (9 for 19XX)
YY = year
JJJ = Julian day

A SIA file and SIA SDB information file exists for each hour that an analysis was performed. Each SIA file has been named using the following convention:

Positions 1-4	Platform:	
		all_ = All satellite platforms are used to create a SIA.
Positions 5-8	Type of file:	
		ian_ = integrated analysis file
		sdb_ = SERCAA data base (SDB) information file
Positions 9-12	Region of interest:	
		(Given in 16th mesh coordinates)
		ban_ = 536,168 600,232
		pan_ = 504,914 568,978
		jap_ = 344,234 408,298
Positions 13-16	Julian day:	
		147_ = Julian day 147 etc. ...
Positions 17-18	Hour:	
		00 = SIA for hour 00 etc. ...
Positions 19-22	Extension:	
		.dat = file extension
Example:		all_ian_ban_147_10.dat

The RE tape contains 243 tar files. Each of these tar files represent a directory that contains all the related data used as input to create at least one of the SIA data files. Each directory name follows the convention:

ENTRY/

where:

ENTRY = the SDB entry number

Each RE file has been named following these guidelines:

Positions 1-4	Platform:	
		n11_ = NOAA N_11
		n12_ = NOAA N_12
		f10_ = DMSP F_10
		f11_ = DMSP F_11
		g07_ = GOES-7
		g04_ = GMS-4
Positions 5-8	Type of file:	
		001_ = satellite data channel 1
		002_ = satellite data channel 2
		...
		...
		005_ = satellite data channel 5
		lat_ = latlon data
		ang_ = angles data

		mcf_ = cloud mask data sdb_ = SDB information file
Positions 9-12	Area of data:	
		sat_ = North America set_ = Asia sdt_ = Africa
Positions 13-16	Julian day:	
		147_ = Julian day 147 etc. ...
Positions 17-18	Hour:	
		00 = hour of the data
Positions 19-22	Extension:	
		.dat = raw data .tif = tif formatted data

Examples:

f10_001_sat_150_14.tif
f10_002_sat_150_14.tif
f10_lat_sat_150_14.tif
f10_ang_sat_150_14.tif
f10_mcf_sat_150_14.tif
f10_sdb_sat_150_14.tif

The exceptions to this guideline are GOES-7 satellite data, GOES-7 SDB information files and missing SDB information files.

GOES-7 files are named using the following guideline:

Positions 1-3	sensor type:	ir1,ir2,ir3 & vis
Positions 4-7	hourminutes:	time in hhmm
Positions 8-11	extension:	.dat = raw data .tif = tif formatted data

Many of the GOES-7 SDB information files were missing and therefore created as part of this release, they are simply named, SDB_entry_number_sdb.dat (i.e. 7199_sdb.dat).

In future releases the GOES-7 satellite data will adhere to the above convention.

Refer to SIA.TAR.LIST and ENTRIES.TAR.LIST for a complete listing of the files contained on the release tapes. Please note, each directory listed represents a different tar file.

****

What are related data items ?
What is the SDB entry number ?
What are related entries ?

The SDB registration process is a process that automatically places descriptive data items about a satellite scan into the SDB. The SDB registration process allocates a group of unique entry numbers to be used as place holders for all of the related data items for a given satellite scan. The related data items consists of satellite, latitude/longitude, angles (Geometry) and product(cloud mask) data. As an example, if a DMSP F_11 scan was to be registered in the SDB, the registration process would request for

a group of five contiguous entry numbers(i.e. 1001-1005). These five entry numbers would be used as place holder for the following related data items:

1001	f11 visible channel
1002	f11 infrared channel
1003	latitude/longitude data
1004	angles(geometry) data
1005	product data

The "SDB entry number" is the first entry number of the group of entry numbers provided by the registration process. The first entry number is used to "key" into the related data items for that group. In the example provided above the SDB entry number would be 1001.

This release process uses the SDB entry number in each group to logically divide the data into separate directories (i.e. the directory name is first SDB entry number for each group of entry numbers). Using the example provided above the directory named "1001/" contains all the related data items for that group (i.e. the directory contains the data for entry 1001 through entry 1005).

To build a SIA it is necessary to use as input, related data items from one or more satellite scans and/or satellite platforms. The SDB entry number is used to keep track of all inputs to the SIA. The list of related entries are given as SDB entry numbers.

****

How do I get a particular SIA data set ?

You must use the UNIX tar utility to extract the data from the tape. By using the tar utility you may extract individual files or the entire directory. Use the SIA.TAR.LIST to determine where to position the tape and then use the appropriate tar command to extract the files you want. For example, if you want to extract all of the SIA data files from the first and second tar files, the following commands might be used:

```
% pwd
/users/smith
% mkdir data
% cd data
% tar xvf /dev/rst8 993147
% tar xvf /dev/rst8 993148
```

Upon completion all of the SIA data for day 147 would reside in directory /users/smith/data/993147 and all the SIA data for day 148 would reside in directory
/users/smith/data/993148.

****

What is the SDB information file ?

The SDB information file is a text file containing selected SDB record items that help describe the actual data. The SIA SDB information file shows what data went into creating the SIA by listing the related entries. The RE SDB information file lists information about the satellite images,

the latlon file, the angles file and the product file(s).

The following is an example SIA SDB information file:

```
[IA]
ZULU_YYJJJ:=93147                                   : Year, Julian day of SIA
ZULU_HH:=23                                         : Hour of SIA
ROI:=PAN                                            : Region of Interest
NUM_RELATED_LAYER:=3                                : Number for related entries
RELATED_LAYER_1:= 4148                              : 1st related SDB entry number
RELATED_LAYER_2:= 7199                              : 2nd related SDB entry number
RELATED_LAYER_3:= 8988                              : 3d related SDB entry number
TDISK:=SDB_Int:
TDIR:=[SERCAA.DATA.993147]
FILE_IA_1:=ALL_IAN_PAN_147_23.Dat                   : SIA file name
SDB_SET:=MAY93                                      : Set identifier May of 1993
```

The following is an example RE SDB information file:

```
[SATIMG]
SAT_CODE:=16                                                  : Satellite code
ZULU_YYJJJ:=93147                                             : Year, Julian day of scan
ZULU_HHMMSS:=82252                                            : Time of scan
NUM_LINES:=1375                                               : Number of lines
ELEM_PER_LINE:=409                                            : Elements per line
BYTES_PER_ELEM:=1                                             : Bytes per element
7199:=AVH$005:[SERCAA.DATA.993147]N11_001_SET_147_08.TIF      : Channel 1 file
7200:=AVH$005:[SERCAA.DATA.993147]N11_002_SET_147_08.TIF      : Channel 2 file
7201:=AVH$005:[SERCAA.DATA.993147]N11_003_SET_147_08.TIF      : Channel 3 file
7202:=AVH$005:[SERCAA.DATA.993147]N11_004_SET_147_08.TIF      : Channel 4 file
7203:=AVH$005:[SERCAA.DATA.993147]N11_005_SET_147_08.TIF      : Channel 5 file

[LATLON]
LL_REC_LEN:=204                                               : Record length in bytes
LL_LINE_INTERVAL:=1                                           : Sub-sample line interval
LL_ELEM_INTERVAL:=8                                           : Sub-sample element interval
LL_ELEM_PER_LINE:=51                                          : Latlon pairs per line
LL_FILE:=AVH$005:[SERCAA.DATA.993147]N11_LAT_SET_147_08.DAT   : latitude/longitude file

[ANGLES]
ANG_REC_LEN:=612                                              : Record length in bytes
ANG_LINE_INTERVAL:=1                                          : Sub-sample line interval
ANG_ELEM_INTERVAL:=8                                          : Sub-sample element interval
ANG_ELEM_PER_LINE:=51                                         : Angles triplets per line
ANG_FILE:=AVH$005:[SERCAA.DATA.993147]N11_ANG_SET_147_08.DAT  : Angles file

[PRODUCT]
7206001:=sdb$prd:[SERCAA.DATA.993147]N11_MCF_SET_147_08.TIF   : Cloud mask file
```

****

How do I know which RE data went into a particular SIA ?

There are two ways to determine which RE data sets went into a particular SIA. The first way is reference the SIA SDB information file. Each "RELATED_LAYER" listed is a reference, by SDB entry number, to the RE data. Use the referred SDB entry number to retrieve the related data from the RE data tape.

For example, refer to the above SIA SDB information file. The "RELATED_LAYERED_1:=4148" line implies that SDB entry number 4148 and the related data items for entry 4148 (along with SDB entry numbers 7199 and 8988)

were used to create "ALL_IAN_PAN_147_23.Dat".

The second way is to read the header information from the SIA file (Please refer to the DATA_DESCRIPTION).

****

How do I get the RE data files ?

Once you have examined the SIA SDB information file and you have identified the related entry numbers, you must use the UNIX tar utility to extract the data from the tape. By using the tar utility you may extract individual files or the entire directory. Use the RE.TAR.LIST to determine where to position the tape and then use the appropriate tar command to extract the files you want. For example, if you want to extract all of the RE data files from the first tar file, the following commands might be used:

```
% pwd
/users/smith
% mkdir data
% cd data
% tar xvf /dev/rst8 7199
```

Upon completion of this command all of the RE data related to SDB entry number 7199 would reside in directory /users/smith/data/7199.

****

For the following question please refer to the example SDB information files as needed.

****

What is the format of the satellite data and how do I access it?

The dimensions of the satellite data are defined by the three parameters, NUM_OF_LINES, ELEM_PER_LINE and BYTES_PER_ELEM . To access the data use the following logic.

If the file extension is ".dat"
then use the appropriate C or FORTRAN read statements.

If the file extension is ".tif"
then use a tiff reader or tiff library (you may view the
images by using the public domain application, XV).

For a detailed explanation, refer to Appendix A.

****

What is the format of the latlon data and how do I access it?

The latlon data are sub-sampled. The dimensions are defined LL_LINE_INTERVAL, LL_ELEM_INTERVAL and LL_ELEM_PER_LINE. LL_ELEM_PER_LINE defines the number of longitude/latitude pairs per line. Each pair is four bytes (two bytes lon, two bytes lat). To access the data use the appropriate C or FORTRAN read statements.

For a detailed explanation, refer to Appendix A.

****
What is the format of the angles data and how do I access it?

The angles data are sub-sampled. The dimensions are defined by ANG_LINE_INTERVAL, ANG_ELEM_INTERVAL and ANG_ELEM_PER_LINE. ANG_ELEM_PER_LINE
defines the number of triplets (satellite-zenith/solar-zenith/azimuth) per line. Each item in the triplet is a float data type. To access the data use the appropriate C or FORTRAN read statements.

For a detailed explanation, refer to Appendix A.

P601

Data Save Documentation Report No. 2

ADVANCED GEOPHYSICAL ENVIRONMENT SIMULATION TECHNIQUES

Task 1: Satellite Data Sets for Worldwide Cloud Prediction

This data documentation report covers data set generation
for the DNA region of interest:

East Asia Area (EASA)

for the period:

22-30 March 1993

Contract Number F19628-94-C-0106

issued by:

Electronic Systems Division
Air Force Systems Command
Hanscom AFB, MA 01731

Submitted by:

Atmospheric and Environmental Research, Inc.
840 Memorial Drive
Cambridge, MA 02139

30 November 1994

David B. Hogan
Gary B. Gustafson
Principal Investigators

1.0 Introduction

This Data Documentation Report provides a description of the second data save made in accordance with the revised statement of work for Satellite Data Sets for Worldwide Cloud Prediction Models. It is intended to provide a description of the data set, its format, how it was gathered and processed, and a description of the algorithms used to generate it. The data set consists of raw satellite data and analyzed products produced by the SERCAA cloud analysis algorithms. The period covered is 22-30 March 1993 for the DNA region of interest: East Asia Area (EASA). This region covers the following (i,j) 16^{th} mesh grid coordinates: 227,13 - 451,395. All available data from those dates are included. These data were processed specifically for DNA using software developed from the SERCAA cloud analysis algorithms described by Gustafson et. al (1994). Substantial modifications were required to the Cloud Layering and Analysis Integration modules to accommodate the high volume of data included in this data set. Two tapes are provided, one with Level 1, 2 and 3 products and the second with Level 4. Data formats for the Level 3 and 4 products differ from those used in the initial May 1993 data set provided earlier (see Data Save Documentation Report No. 1, dated 14 October 1994).

2.0 Processing Environment

Satellite data processing for this data set used the SERCAA cloud analysis algorithms described by Gustafson et al. (1994). Multisource data from the DMSP F10 and F11, NOAA-11 and NOAA-12, and GMS-4 satellites were used. Data sources were as follows: DMSP - National Geophysical Data Center (NGDC), Boulder, CO; NOAA - National Climatic Data Center (NCDC), Ashville, NC; GMS - Sea Space Corp., San Diego, CA. All data were obtained by the Phillips Laboratory and were received on tape in various formats. All data processing was performed on the Air Force Interactive Meteorological System (AIMS) at the Phillips Laboratory. The SERCAA cloud analysis algorithms use four levels of data processing as summarized in Figure 1.

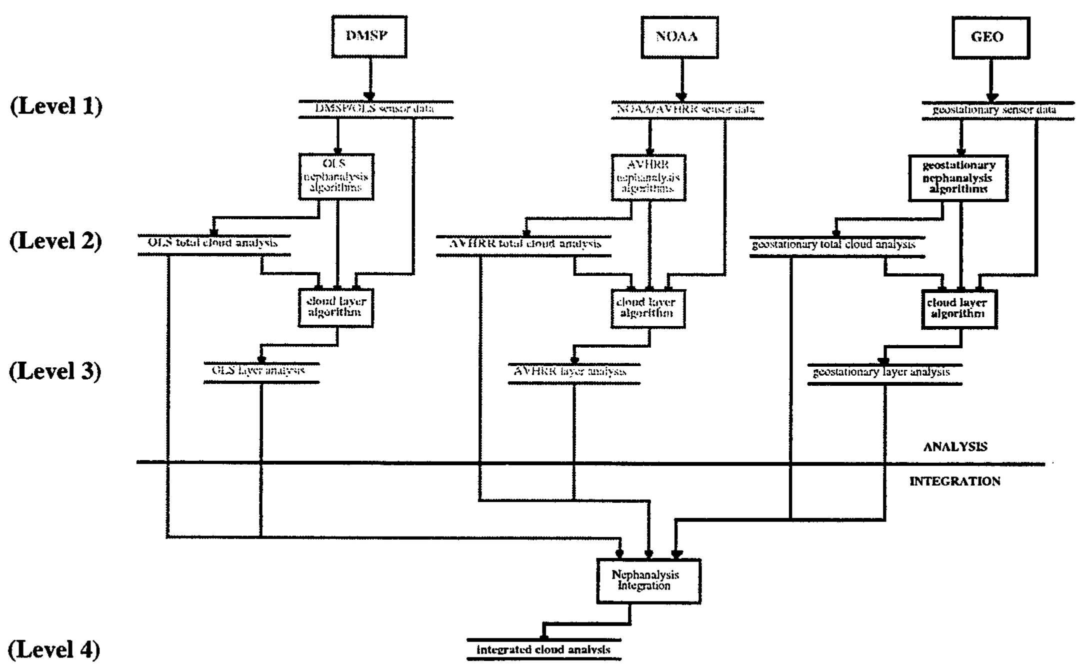


Figure 1 SERCAA data flow and processing levels

Level 1 processing consists of data ingest. Tape data are processed through separate ingest programs depending on the data source and format. All data are then stored in a standard format in the original satellite scan projection. The format consists of flat files where the number of elements correspond to the number of pixels in the satellite scan line and the number of rows corresponds to the number of scan lines. Data are maintained on AIMS through the SERCAA Database (SDB) management software. Level 1 data products consist of separate files for each sensor channel plus two additional files containing Earth location and satellite/solar geometry information. Satellite data characteristics are summarized in Table 1. In cases where visible and infrared channel resolution differ, the higher resolution data are subsampled to match the coarser resolution data (e.g., GMS visible data are subsampled by a factor of four to match the IR data resolution). Earth location data consist of latitude-longitude pairs that are maintained at a subsampled resolution relative to the satellite data. For each sensor scan line, one latitude-longitude pair is provided for every n^{th} pixel, where n varies with satellite. Geometry information are also subsampled in the same ratio as the Earth location information and consist of three angles: satellite zenith, solar zenith, and sun-satellite azimuth. Ingest products are described more completely in Section 2 of Gustafson et al. (1994).

Table 1. Sensor Channel Data Attributes During SERCAA

Satellite	Sensor	Channel (μm)	Data Format	Resolution[1] (km)	Bits per Pixel[2]	Pixels per Scan Line
DMSP	OLS	0.40-1.10	counts	2.7	6	1464
		10.5-12.6	EBBT	2.7	8	1464
NOAA	AVHRR	0.58-0.68	percent albedo	4.0	10	409
		0.72-1.10	percent albedo	4.0	10	409
		3.55-3.93	EBBT	4.0	10	409
		10.3-11.3	EBBT	4.0	10	409
		11.5-12.5	EBBT	4.0	10	409
GMS	VISSR	0.5-0.75	counts	1.25	6	10000
		10.5-12.5	EBBT	5.0	8	2500

[1]Sensor resolution at satellite subpoint that will provide global coverage.

[2]AVHRR radiance data are transmitted at 10-bit resolution, however, the SERCAA development system could only accommodate 8-bit brightness temperature data (although the full 10-bit resolution is used in the radiance-to-brightness-temperature transformation).

Level 2 processing consists of sensor-specific nephanalysis algorithms. Level 1 sensor data from DMSP, NOAA, and the GMS geostationary satellites are processed through separate algorithms as indicated in Figure 1. Each time data from a new satellite pass are ingested, they are analyzed through the appropriate nephanalysis algorithm and results are placed in a Level 2 output file. One output file is generated for each nephanalysis run and nephanalysis results are stored in the original satellite scan projection with one byte of information for each pixel. Each byte is bit-packed according to the map in Table 2. For each set of Level 1 products generated from a satellite pass, one Level 2 product file is generated.

Table 2. Cloud Analysis Algorithm MCF File Bit Assignments

Bit	Assignment	Description
0	Cloud Mask	ON = Cloud-Filled OFF = Cloud-Free
1	Low Cloud	ON = Low Cloud Found
2	Thin Cirrus Cloud	ON = Thin Cirrus Cloud Found
3	Precipitating Cloud	ON = Precipitating Cloud Found
4	Partial Cloud	Only used by DMSP algorithm
5	Data Dropout	ON = Missing or Unreliable Data
6	Confidence	0 = Missing Data; 1 = Low;
7	Flag	2 = Middle; 3 = High

Level 3 processing uses Level 1 and 2 products as input to segment the cloudy regions into vertical cloud layers and to classify different cloud types. It also remaps the data from the individual satellite projections to the AFGWC standard polar stereographic map projection (Hoke et al., 1981) at 16^{th} mesh grid resolution. The EASA region of interest processed for the March 1993 data set have the following (i,j) 16^{th} mesh grid coordinates: $395 \leq i \leq 451$, $13 \leq j \leq 227$. Level 3 products are generated for each 16^{th} mesh grid cell and contain the information in Table 3. A maximum of four cloud layers can be identified for each grid cell. One Level 3 file is created for each set of Level 1 and 2 products. All Level 1, 2, and 3 products associated with a single satellite pass are related through SDB and are provided on the DNA tapes as a set. Note that for the EASA region, all Level 3 files are a fixed size of 225x383 grid cells.

Table 3. Cloud Typing and Layering Output

Parameter	Description
i	16^{th} mesh i coordinate for Grid Cell
j	16^{th} mesh j coordinate for Grid Cell
sdb_ir_entry	SDB entry number of input IR sensor data
ddd	Sensor data Julian date
hhmm	Sensor data valid time (UTC)
layer_var(4)	Cloud top IR variance of pixels in each layer
meantemp(4)	Cloud top mean IR Temperature of pixels in each layer
n_layer_pix(4)	Total number of pixels in each layer
cloud_type(4)	Cloud type of each layer
low_cloud(4)	Number of low cloud pixels in this layer detected by cloud analysis algorithm
thin_cirrus(4)	Number of thin cirrus pixels in this layer detected by cloud analysis algorithm
precip(4)	Number of precipitating cloud pixels in this layer detected by cloud analysis algorithm
sunrise	Local sunrise time (UTC)
sunset	Local sunset time (UTC)
vid	Satellite vehicle (platform) ID
num_pixels	Total number of satellite pixels in 16^{th} mesh grid cell
dropouts	Number of bad data pixels in 16^{th} mesh grid cell
partial	Number of partial cloud pixels detected by DMSP cloud analysis algorithm

Level 4 processing is a clock driven process with one new Level 4 integrated analysis performed each hour. Thus, integration is differentiated from the Level 1, 2, and 3 products are event-driven (i.e., resulting from the ingest of a new satellite pass). The integration module operates on the most recent Level 3 gridded products available from each satellite source. Like Level 3 products, the Level 4 output files also conform to the AFGWC 16th mesh grid structure; output parameters for each grid cell are summarized in Table 4.

Table 4 Analysis Integration Processed Parameters

Parameter	Description
i	16th mesh i (column) coordinate
j	16th mesh j (row) coordinate
nlay	Number of Cloud Layers
cftot	Total Cloud Fraction
cf(4)	Layer Cloud Fraction
ctt(4)	Layer Cloud Top IR Temperature (K)
ctz(4)	Layer clout top height (m)
ity(4)	Layer Cloud Type
ecft	Estimated Error in Total Cloud Fraction
ecf(4)	Estimated Error in Layer Cloud Fraction
icf(4)	Analysis Confidence Flag Index For Each Layer
sdb(3)	SDB entry number of input analyses (NOAA, DMSP, GMS)

3.0 Tape Format

All data for the March 1993 EASA data save are contained on two 8 mm tapes written in UNIX tar format. The first tape, labeled: DNA MAR93 ENTRIES, contains all the Level 1-3 products. The second tape, labeled: DNA MAR93 IA, contains all Level 4 products. The size of the combined Level 1, 2 and 3 products is approximately 1.9 Gbytes and the Level 4 products occupy 1.2 Gbytes. In addition to the two tapes, hard-copy listings of the contents of the Level 4 tape are also provided. The corresponding listing of the Level 1-3 tape is very large, so a UNIX script is provided to generate a listing at the user's site. It may be useful to place the listing file generated by the script into an edit program to scan and search it quickly. The listings are required to locate specific data sets on the tapes.

Level 1-3 products are generated for each new pass of satellite data received during the period of the data save. Appendix A contains a chronological list of each satellite pass used to produce the March 93 data sets. All available data for the period covered were included; any gaps in the data list are due to either missing or bad data. Numerous DMSP orbits contained periodic data dropouts as illustrated in Figure 2, the most severely affected files were removed from the data set. For data archiving purposes all Level 1-3 products associated with a given satellite pass were placed in a single directory and subsequently placed on tape as a single tar file. Thus the first tape contains a series of several hundred tar files; each file contains all Level 1-3 products associated with a single satellite pass. Level 4 files are grouped on the second tape by day, thus for the March data save there are nine tar files on the Level 4 tape that each contains all Level 4 output files for each of the nine days 93081-93089 (22-30 March 1993). For each set of

Level 1-3 products, and for each Level 4 file there is also an SDB Information File. These files contain descriptive metadata information extracted from the SERCAA Database that describe the relevant attributes of the SERCAA product files. For example, information files list the number of pixels in a scan line of satellite data and the number of scan lines in the file. Information on subsampling ratios for the Earth location and angles files are also contained there.

Figure 2 Sample of DMSP periodic data dropouts.

Detailed descriptions of the file formats used for each output level, and the associated information files, provided for the May 1993 save (Level 1, 2, 3, and 4) are provided in Appendix B. Appendix C provides a guide for extracting data sets from tape.

4.0 References

Gustafson, G.B., R.G. Isaacs, R.P. d'Entremont, J.M. Sparrow, T.M. Hamill, C. Grassotti, D.W. Johnson, C.P. Sarkisian, D.C. Peduzzi, B.T. Pearson, V.D. Jakabhazy, J.S. Belfiore, and A.S. Lisa, 1994: Support of Environmental Requirements for Cloud Analysis and Archive (SERCAA): algorithm descriptions. PL-TR-94-2114, Phillips Laboratory, Hanscom AFB, MA, ADA283240.

Hoke, J.E., J.L. Hayes, L.G. Renninger, 1981: Map projections and grid systems for meteorological applications. AFGWC-TN-79-003, Air Weather Service, Scott, AFB, IL.

Appendix A

Chronological List of Input Satellite Data

Entry	Satellite	Date	Time	ROI	Cols	Rows	Resolution
1	GMS 4	93081	013300	EAS	1376	1010	5.00
16	GMS 4	93081	023300	EAS	1376	1010	5.00
21	GMS 4	93081	033300	EAS	1376	1010	5.00
26	GMS 4	93081	042600	EAS	1376	1010	5.00
948	NOAA 11	93081	050222	EAS	409	1072	4.00
11	GMS 4	93081	053300	EAS	1376	1010	5.00
36	GMS 4	93081	063400	EAS	1376	1010	5.00
940	NOAA 11	93081	063841	EAS	409	1714	4.00
1769	DMSP F11	93081	073000	EAS	1465	1010	2.70
41	GMS 4	93081	073400	EAS	1376	1010	5.00
1774	DMSP F11	93081	080700	EAS	1465	2023	2.70
46	GMS 4	93081	083400	EAS	1376	1010	5.00
51	GMS 4	93081	093400	EAS	1376	1010	5.00
1212	NOAA 12	93081	101328	EAS	409	1698	4.00
56	GMS 4	93081	102700	EAS	1376	1010	5.00
1779	DMSP F11	93081	104700	EAS	1465	774	2.70
1629	DMSP F10	93081	112200	EAS	1465	1890	2.70
61	GMS 4	93081	113300	EAS	1376	1010	5.00
1204	NOAA 12	93081	115314	EAS	409	667	4.00
66	GMS 4	93081	123400	EAS	1376	1010	5.00
1634	DMSP F10	93081	135900	EAS	1465	1419	2.70
71	GMS 4	93081	154400	EAS	1376	1010	5.00
76	GMS 4	93081	162700	EAS	1376	1010	5.00
964	NOAA 11	93081	173007	EAS	409	626	4.00
81	GMS 4	93081	173300	EAS	1376	1010	5.00
86	GMS 4	93081	183300	EAS	1376	1010	5.00
956	NOAA 11	93081	190724	EAS	409	1887	4.00
91	GMS 4	93081	193300	EAS	1376	1010	5.00
96	GMS 4	93081	203300	EAS	1376	1010	5.00
101	GMS 4	93081	213300	EAS	1376	1010	5.00
106	GMS 4	93081	222600	EAS	1376	1010	5.00
1220	NOAA 12	93081	223638	EAS	409	1793	4.00
111	GMS 4	93081	233200	EAS	1376	1010	5.00
1639	DMSP F10	93082	003900	EAS	1465	1572	2.70
121	GMS 4	93082	013300	EAS	1376	1010	5.00
126	GMS 4	93082	023300	EAS	1376	1010	5.00
131	GMS 4	93082	033400	EAS	1376	1010	5.00
136	GMS 4	93082	042600	EAS	1376	1010	5.00
980	NOAA 11	93082	045031	EAS	409	901	4.00
141	GMS 4	93082	053300	EAS	1376	1010	5.00
988	NOAA 11	93082	062717	EAS	409	1700	4.00
146	GMS 4	93082	063400	EAS	1376	1010	5.00
151	GMS 4	93082	073400	EAS	1376	1010	5.00
972	NOAA 11	93082	080759	EAS	409	566	4.00
156	GMS 4	93082	083400	EAS	1376	1010	5.00
161	GMS 4	93082	093400	EAS	1376	1010	5.00
166	GMS 4	93082	102700	EAS	1376	1010	5.00
1308	NOAA 12	93082	113132	EAS	409	1049	4.00
171	GMS 4	93082	113300	EAS	1376	1010	5.00
1644	DMSP F10	93082	115200	EAS	1465	1520	2.70
1649	DMSP F10	93082	122700	EAS	1465	1850	2.70

Entry	Satellite	Date	Time	ROI	Cols	Rows	Resolution
176	GMS 4	93082	123400	EAS	1376	1010	5.00
181	GMS 4	93082	154400	EAS	1376	1010	5.00
186	GMS 4	93082	162700	EAS	1376	1010	5.00
191	GMS 4	93082	173300	EAS	1376	1010	5.00
196	GMS 4	93082	183300	EAS	1376	1010	5.00
1428	NOAA 11	93082	185529	EAS	409	1841	4.00
201	GMS 4	93082	193300	EAS	1376	1010	5.00
206	GMS 4	93082	203300	EAS	1376	1010	5.00
211	GMS 4	93082	213300	EAS	1376	1010	5.00
1316	NOAA 12	93082	221545	EAS	409	1637	4.00
216	GMS 4	93082	222600	EAS	1376	1010	5.00
221	GMS 4	93082	233300	EAS	1376	1010	5.00
1324	NOAA 12	93082	235415	EAS	409	2033	4.00
231	GMS 4	93083	013300	EAS	1376	1010	5.00
1654	DMSP F10	93083	014500	EAS	1465	2262	2.70
236	GMS 4	93083	023300	EAS	1376	1010	5.00
241	GMS 4	93083	033300	EAS	1376	1010	5.00
246	GMS 4	93083	042600	EAS	1376	1010	5.00
1236	NOAA 11	93083	043837	EAS	409	691	4.00
251	GMS 4	93083	053300	EAS	1376	1010	5.00
1244	NOAA 11	93083	061549	EAS	409	1682	4.00
256	GMS 4	93083	063400	EAS	1376	1010	5.00
261	GMS 4	93083	073400	EAS	1376	1010	5.00
1228	NOAA 11	93083	075541	EAS	409	804	4.00
266	GMS 4	93083	083400	EAS	1376	1010	5.00
1340	NOAA 12	93083	093240	EAS	409	1503	4.00
271	GMS 4	93083	093400	EAS	1376	1010	5.00
276	GMS 4	93083	102700	EAS	1376	1010	5.00
1332	NOAA 12	93083	110950	EAS	409	1352	4.00
281	GMS 4	93083	113400	EAS	1376	1010	5.00
286	GMS 4	93083	123400	EAS	1376	1010	5.00
1659	DMSP F10	93083	125700	EAS	1465	1991	2.70
1664	DMSP F10	93083	143600	EAS	1465	679	2.70
291	GMS 4	93083	154400	EAS	1376	1010	5.00
296	GMS 4	93083	162700	EAS	1376	1010	5.00
301	GMS 4	93083	173300	EAS	1376	1010	5.00
1252	NOAA 11	93083	184337	EAS	409	1789	4.00
306	GMS 4	93083	193600	EAS	1376	1010	5.00
1799	DMSP F11	93083	200900	EAS	1465	1976	2.70
1260	NOAA 11	93083	202635	EAS	409	946	4.00
311	GMS 4	93083	203600	EAS	1376	1010	5.00
316	GMS 4	93083	213600	EAS	1376	1010	5.00
1348	NOAA 12	93083	215501	EAS	409	1220	4.00
321	GMS 4	93083	222900	EAS	1376	1010	5.00
1804	DMSP F11	93083	224800	EAS	1465	2442	2.70
1356	NOAA 12	93083	233258	EAS	409	1983	4.00
1669	DMSP F10	93084	001500	EAS	1465	2129	2.70
1268	NOAA 11	93084	060417	EAS	409	1519	4.00
1809	DMSP F11	93084	083000	EAS	1465	1870	2.70
1814	DMSP F11	93084	090800	EAS	1465	1534	2.70
1372	NOAA 12	93084	091201	EAS	409	1279	4.00

Entry	Satellite	Date	Time	ROI	Cols	Rows	Resolution
1364	NOAA 12	93084	104808	EAS	409	1591	4.00
1674	DMSP F10	93084	112700	EAS	1465	1927	2.70
1679	DMSP F10	93084	130500	EAS	1465	1323	2.70
331	GMS 4	93084	162800	EAS	1376	1010	5.00
336	GMS 4	93084	173400	EAS	1376	1010	5.00
1276	NOAA 11	93084	183147	EAS	409	1732	4.00
376	GMS 4	93084	183500	EAS	1376	1010	5.00
381	GMS 4	93084	193600	EAS	1376	1010	5.00
1284	NOAA 11	93084	201059	EAS	409	2066	4.00
386	GMS 4	93084	203600	EAS	1376	1010	5.00
1380	NOAA 12	93084	213429	EAS	409	867	4.00
391	GMS 4	93084	213600	EAS	1376	1010	5.00
396	GMS 4	93084	222900	EAS	1376	1010	5.00
1388	NOAA 12	93084	231146	EAS	409	1920	4.00
401	GMS 4	93084	233500	EAS	1376	1010	5.00
1684	DMSP F10	93085	012200	EAS	1465	2386	2.70
346	GMS 4	93085	013400	EAS	1376	1010	5.00
351	GMS 4	93085	023400	EAS	1376	1010	5.00
356	GMS 4	93085	033400	EAS	1376	1010	5.00
361	GMS 4	93085	042600	EAS	1376	1010	5.00
366	GMS 4	93085	053300	EAS	1376	1010	5.00
1300	NOAA 11	93085	055241	EAS	409	1632	4.00
1292	NOAA 11	93085	073104	EAS	409	1201	4.00
406	GMS 4	93085	073300	EAS	1376	1010	5.00
411	GMS 4	93085	083300	EAS	1376	1010	5.00
1396	NOAA 12	93085	085112	EAS	409	1020	4.00
416	GMS 4	93085	093300	EAS	1376	1010	5.00
1819	DMSP F11	93085	095500	EAS	1465	1725	2.70
421	GMS 4	93085	102600	EAS	1376	1010	5.00
1404	NOAA 12	93085	102657	EAS	409	1717	4.00
426	GMS 4	93085	113300	EAS	1376	1010	5.00
431	GMS 4	93085	123300	EAS	1376	1010	5.00
436	GMS 4	93085	154300	EAS	1376	1010	5.00
441	GMS 4	93085	162600	EAS	1376	1010	5.00
446	GMS 4	93085	173300	EAS	1376	1010	5.00
996	NOAA 11	93085	182000	EAS	409	1502	4.00
451	GMS 4	93085	183300	EAS	1376	1010	5.00
456	GMS 4	93085	193300	EAS	1376	1010	5.00
1004	NOAA 11	93085	195854	EAS	409	2038	4.00
461	GMS 4	93085	203300	EAS	1376	1010	5.00
1824	DMSP F11	93085	204400	EAS	1465	1374	2.70
1412	NOAA 12	93085	211411	EAS	409	544	4.00
466	GMS 4	93085	213300	EAS	1376	1010	5.00
471	GMS 4	93085	222600	EAS	1376	1010	5.00
1420	NOAA 12	93085	225040	EAS	409	1849	4.00
1689	DMSP F10	93085	231400	EAS	1465	1063	2.70
476	GMS 4	93085	233300	EAS	1376	1010	5.00
486	GMS 4	93086	013300	EAS	1376	1010	5.00
491	GMS 4	93086	023300	EAS	1376	1010	5.00
496	GMS 4	93086	033300	EAS	1376	1010	5.00
501	GMS 4	93086	042600	EAS	1376	1010	5.00

Entry	Satellite	Date	Time	ROI	Cols	Rows	Resolution
506	GMS 4	93086	053300	EAS	1376	1010	5.00
1020	NOAA 11	93086	054103	EAS	409	1510	4.00
511	GMS 4	93086	063300	EAS	1376	1010	5.00
1829	DMSP F11	93086	070500	EAS	1465	1555	2.70
1012	NOAA 11	93086	071845	EAS	409	1363	4.00
516	GMS 4	93086	073300	EAS	1376	1010	5.00
1452	NOAA 12	93086	083014	EAS	409	689	4.00
521	GMS 4	93086	083400	EAS	1376	1010	5.00
526	GMS 4	93086	093300	EAS	1376	1010	5.00
1834	DMSP F11	93086	094200	EAS	1465	1890	2.70
1444	NOAA 12	93086	100645	EAS	409	1686	4.00
531	GMS 4	93086	102600	EAS	1376	1010	5.00
1694	DMSP F10	93086	102700	EAS	1465	1163	2.70
536	GMS 4	93086	113300	EAS	1376	1010	5.00
1436	NOAA 12	93086	114602	EAS	409	801	4.00
1699	DMSP F10	93086	120200	EAS	1465	1998	2.70
541	GMS 4	93086	123300	EAS	1376	1010	5.00
1704	DMSP F10	93086	144200	EAS	1465	542	2.70
546	GMS 4	93086	154300	EAS	1376	1010	5.00
551	GMS 4	93086	162600	EAS	1376	1010	5.00
556	GMS 4	93086	173300	EAS	1376	1010	5.00
1028	NOAA 11	93086	180816	EAS	409	1264	4.00
561	GMS 4	93086	183300	EAS	1376	1010	5.00
566	GMS 4	93086	193300	EAS	1376	1010	5.00
1036	NOAA 11	93086	194651	EAS	409	2007	4.00
1839	DMSP F11	93086	203200	EAS	1465	1118	2.70
571	GMS 4	93086	203300	EAS	1376	1010	5.00
576	GMS 4	93086	213300	EAS	1376	1010	5.00
581	GMS 4	93086	222600	EAS	1376	1010	5.00
1460	NOAA 12	93086	222941	EAS	409	1759	4.00
586	GMS 4	93086	233300	EAS	1376	1010	5.00
1468	NOAA 12	93087	001049	EAS	409	1557	4.00
596	GMS 4	93087	013300	EAS	1376	1010	5.00
601	GMS 4	93087	023300	EAS	1376	1010	5.00
606	GMS 4	93087	033300	EAS	1376	1010	5.00
611	GMS 4	93087	042600	EAS	1376	1010	5.00
1052	NOAA 11	93087	052921	EAS	409	1384	4.00
616	GMS 4	93087	053300	EAS	1376	1010	5.00
621	GMS 4	93087	063300	EAS	1376	1010	5.00
1044	NOAA 11	93087	070627	EAS	409	1504	4.00
626	GMS 4	93087	073300	EAS	1376	1010	5.00
1844	DMSP F11	93087	075300	EAS	1465	1380	2.70
1849	DMSP F11	93087	082900	EAS	1465	2033	2.70
631	GMS 4	93087	083300	EAS	1376	1010	5.00
636	GMS 4	93087	093300	EAS	1376	1010	5.00
1484	NOAA 12	93087	094623	EAS	409	1643	4.00
641	GMS 4	93087	102600	EAS	1376	1010	5.00
1709	DMSP F10	93087	105600	EAS	1465	509	2.70
1476	NOAA 12	93087	112420	EAS	409	1156	4.00
646	GMS 4	93087	113300	EAS	1376	1010	5.00
651	GMS 4	93087	123300	EAS	1376	1010	5.00

Entry	Satellite	Date	Time	ROI	Cols	Rows	Resolution
1714	DMSP F10	93087	123300	EAS	1465	1939	2.70
1719	DMSP F10	93087	131100	EAS	1465	1224	2.70
656	GMS 4	93087	154300	EAS	1376	1010	5.00
661	GMS 4	93087	162600	EAS	1376	1010	5.00
666	GMS 4	93087	173300	EAS	1376	1010	5.00
1060	NOAA 11	93087	175636	EAS	409	1050	4.00
671	GMS 4	93087	183300	EAS	1376	1010	5.00
676	GMS 4	93087	193300	EAS	1376	1010	5.00
1068	NOAA 11	93087	193450	EAS	409	1973	4.00
681	GMS 4	93087	203300	EAS	1376	1010	5.00
686	GMS 4	93087	213300	EAS	1376	1010	5.00
1492	NOAA 12	93087	220851	EAS	409	1489	4.00
691	GMS 4	93087	222600	EAS	1376	1010	5.00
696	GMS 4	93087	233300	EAS	1376	1010	5.00
1500	NOAA 12	93087	234710	EAS	409	2018	4.00
1724	DMSP F10	93088	005000	EAS	1465	1839	2.70
1729	DMSP F10	93088	012800	EAS	1465	2406	2.70
706	GMS 4	93088	013300	EAS	1376	1010	5.00
711	GMS 4	93088	023300	EAS	1376	1010	5.00
716	GMS 4	93088	033300	EAS	1376	1010	5.00
721	GMS 4	93088	042600	EAS	1376	1010	5.00
1076	NOAA 11	93088	051736	EAS	409	1250	4.00
726	GMS 4	93088	053300	EAS	1376	1010	5.00
731	GMS 4	93088	063300	EAS	1376	1010	5.00
1084	NOAA 11	93088	065408	EAS	409	1629	4.00
736	GMS 4	93088	073300	EAS	1376	1010	5.00
1604	DMSP F11	93088	074000	EAS	1465	1183	2.70
1609	DMSP F11	93088	081700	EAS	1465	2042	2.70
741	GMS 4	93088	083300	EAS	1376	1010	5.00
1516	NOAA 12	93088	092550	EAS	409	1430	4.00
746	GMS 4	93088	093300	EAS	1376	1010	5.00
751	GMS 4	93088	102600	EAS	1376	1010	5.00
1624	DMSP F11	93088	105800	EAS	1465	534	2.70
1508	NOAA 12	93088	110238	EAS	409	1436	4.00
1734	DMSP F10	93088	110300	EAS	1465	1654	2.70
756	GMS 4	93088	113300	EAS	1376	1010	5.00
761	GMS 4	93088	123300	EAS	1376	1010	5.00
1739	DMSP F10	93088	133900	EAS	1465	1712	2.70
766	GMS 4	93088	154300	EAS	1376	1010	5.00
771	GMS 4	93088	162600	EAS	1376	1010	5.00
776	GMS 4	93088	173300	EAS	1376	1010	5.00
1092	NOAA 11	93088	174500	EAS	409	851	4.00
781	GMS 4	93088	183300	EAS	1376	1010	5.00
1784	DMSP F11	93088	190800	EAS	1465	645	2.70
1100	NOAA 11	93088	192250	EAS	409	1934	4.00
786	GMS 4	93088	193300	EAS	1376	1010	5.00
791	GMS 4	93088	203300	EAS	1376	1010	5.00
796	GMS 4	93088	213300	EAS	1376	1010	5.00
1524	NOAA 12	93088	214810	EAS	409	1099	4.00
801	GMS 4	93088	222600	EAS	1376	1010	5.00
1532	NOAA 12	93088	232554	EAS	409	1965	4.00

Entry	Satellite	Date	Time	ROI	Cols	Rows	Resolution
806	GMS 4	93088	233300	EAS	1376	1010	5.00
816	GMS 4	93089	013300	EAS	1376	1010	5.00
1744	DMSP F10	93089	015700	EAS	1465	2312	2.70
821	GMS 4	93089	023300	EAS	1376	1010	5.00
826	GMS 4	93089	033300	EAS	1376	1010	5.00
831	GMS 4	93089	042600	EAS	1376	1010	5.00
1116	NOAA 11	93089	050548	EAS	409	1104	4.00
836	GMS 4	93089	053300	EAS	1376	1010	5.00
841	GMS 4	93089	063300	EAS	1376	1010	5.00
1108	NOAA 11	93089	064201	EAS	409	1718	4.00
846	GMS 4	93089	073300	EAS	1376	1010	5.00
851	GMS 4	93089	083300	EAS	1376	1010	5.00
1540	NOAA 12	93089	090507	EAS	409	1198	4.00
856	GMS 4	93089	093300	EAS	1376	1010	5.00
861	GMS 4	93089	102600	EAS	1376	1010	5.00
1548	NOAA 12	93089	104056	EAS	409	1647	4.00
1749	DMSP F10	93089	113300	EAS	1465	1247	2.70
1754	DMSP F10	93089	120800	EAS	1465	2007	2.70
871	GMS 4	93089	154300	EAS	1376	1010	5.00
1124	NOAA 11	93089	173329	EAS	409	662	4.00
1132	NOAA 11	93089	191052	EAS	409	1888	4.00
1789	DMSP F11	93089	195600	EAS	1465	417	2.70
1556	NOAA 12	93089	212742	EAS	409	759	4.00
1564	NOAA 12	93089	230444	EAS	409	1901	4.00

Appendix B

Archive Data Format Descriptions

By Data Processing Level

Level 1: Satellite Image Files

Satellite image filenames as they appear on tape have the following naming convention:

SSS_CCC_ROI_DDD_HH.Tif

where

SSS - Satellite identifier:

F10	DMSP F-10
F11	DMSP F-11
N11	NOAA-11
N12	NOAA-12
G04	GMS (Japan)

CCC - spectral channel identifier

ROI - Region of Interest:

EAS for East Asia Area

DDD - Julian day for which the image data are valid

HH - UTC hour of image data

Tif - TIFF file format

File and Record Structure

All image files contain fixed-length records. The number of lines and number of elements in an image file are contained in the Related Entries (RE) SDB information file that is provided with the tape, under the heading of SATIMG:

NUM_LINES	Number of image data lines in the file.
ELEM_PER_LINE	Number of elements (pixels) per line.
BYTES_PER_ELEMENT	Number of bytes per pixel. This number is 1 for all SERCAA imager sensor data.

Image file data are stored in Tagged Image File Format (TIFF), therefore an alternative way to determine image dimensions is to read the TIFF header and examine the width and height fields.

Image pixel values represent either counts or albedo for visible data, and brightness temperatures for thermal infrared data. Table B-1 summarizes the attributes of the SERCAA image data values.

Table B-1 Satellite image characteristics

Satellite ID (SSS)	Spectral Channel (CCC)	Channel Type	Wavelength Band	Physical Value
F10 or F11	001	Visible	0.4 - 1.1 μm	Counts[1]
	002	Long-Wave IR	10 -12 μm	Brightness Temp.[2]
N11 or N12	001	Visible	0.63 μm	Albedo[3]
	002	Near-IR	0.86 μm	Albedo
	003	Mid-Wave IR	3.7 μm	Brightness Temp.
	004	Long-Wave IR	10.7 μm	Brightness Temp.
	005	Long-Wave IR	11.8 μm	Brightness Temp.
G04	001	Visible	0.55 - 0.75 μm	Counts
	002	Long-Wave IR	10.2 - 11.2 μm	Brightness Temp.

[1]Visible counts range from 0 - 255. High counts denote highly reflective surfaces and low counts denote poorly reflective surfaces.

[2]Brightness temperatures are byte-encoded such that the range 0 - 255 corresponds to the temperature range 327.5 K to 200.0 K. The relation between byte values and temperature is linear over this range; the conversion from byte value B to brightness temperature T is given by the relation:

$$T = -0.5\,B + 327.5.$$

[3]Albedo values are byte-encoded such that the range 0 - 255 corresponds to the albedo range 0 - 100%. The relation between byte values and percent albedo is linear; the conversion from byte value B to percent albedo A is given by the relation

$$A = 0.392\,B.$$

Level 1: Latitude-Longitude File

Latitude-longitude filenames as they appear on tape have the following naming convention:

SSS_LAT_ROI_DDD_HH.Dat

where

SSS - Satellite identifier:

F10	DMSP F-10
F11	DMSP F-11
N11	NOAA-11
N12	NOAA-12
G04	GMS (Japan)

LAT - a constant that identifies the file as a latitude-longitude file
ROI - Region of Interest for which the latitude-longitude file is valid:
EAS for East Asia Area
DDD - Julian day of satellite data for which the Earth locations are valid
HH - UTC hour of the satellite data for which the Earth locations are valid

File and Record Structure

Latitude-longitude Earth location files contain fixed-length records, the number and size of which depend on both the size of the corresponding image files and the satellite type. There is always one latitude-longitude record corresponding to each satellite image file record, where a satellite image file record contains one image scan line of information. However, latitude-longitude data are subsampled, relative to the sensor data, along a scan line. There is one latitude-longitude pair for every n^{th} image pixel, where n is a function of satellite. A linear interpolation is used to compute Earth location for intermediate pixels between latitude-longitude reference points.

The information necessary for interpreting a latitude-longitude file record is contained in the Related Entries (RE) SDB information file provided with the tape, under the heading of LATLON:

LL_REC_LEN	Record length in bytes.
LL_LINE_INTERVAL	The number of image file records per lat-lon record. For the March 1993 data set this number is always 1.
LL_ELEM_INTERVAL	The subsampling rate of lat-lon information relative to the corresponding satellite data. For example, if LL_ELEM_INTERVAL = 40, there is one latitude-longitude pair for every 40th image pixel in the scan line (i.e., for pixels 1, 41, 81, ...). Linear interpolation is required to retrieve Earth location information for intermediate pixels 2-40, 42-80, ...
LL_ELEM_PER_LINE	This is the number of latitude-longitude elements per latitude-longitude file record.

A latitude-longitude file data element is a 4-byte structure that contains the scaled latitude and longitude for a given pixel. Thus the length of a latitude-longitude file record in bytes is given by:

LL_REC_LEN = 4 * LL_ELEM_PER_LINE

The 4 bytes consist of two 16-bit integer variables: LONG and LAT. The storage convention is as follows:

LONG	Pixel longitude * 128. To obtain the floating-point longitude, FLONG = LONG / 128. Longitude range is -180° to 180°, positive east.
LAT	Pixel latitude * 128. to obtain floating-point latitude, FLAT = LAT / 128. Latitude range is -90° to 90°, positive north.

Level 1: Angles File

The angles filenames as they appear on tape have the following naming convention:

SSS_ANG_ROI_DDD_HH.Dat

where

SSS - Satellite identifier:

F10	DMSP F-10
F11	DMSP F-11
N11	NOAA-11
N12	NOAA-12
G04	GMS-4 (Japan)

ANG - a constant that identifies the file as an angles file
ROI - Region of Interest for which the angles file is valid:
EAS for East Asia Area
DDD - Julian day of satellite data for which the angles are valid
HH - UTC hour of the satellite data for which the angles are valid

File and Record Structure

Angle files contain fixed-length records, the number and size of which depend on both the size of the corresponding image files and the satellite type. There is always one angles record corresponding to each satellite image file record, where a satellite image file record contains one image scan line of information. However, angle data are subsampled, relative to the sensor data, along a scan line. There is one set of angles for every n^{th} image pixel, where n is a function of satellite. A linear interpolation is used to compute angle values for intermediate pixels between angle reference points.

The information necessary for interpreting an angles file record is contained in the Related Entries (RE) SDB information file provided with the tape, under the heading of ANGLES:

ANG_REC_LEN	Record length in bytes.
ANG_LINE_INTERVAL	The number of image file records per angles record. This number is almost always 1.
ANG_ELEM_INTERVAL	The subsampling rate of angles information relative to the corresponding satellite image. For example, if ANG_ELEM_INTERVAL = 8, there is one set of angles valid for every eighth image pixel in the scan line (i.e., for pixels 1, 9, 17, 25, ...). Linear interpolation is required to retrieve angles information for intermediate pixels 2-8, 10-16, 18-24, ...
ANG_ELEM_PER_LINE	This is the number of angles elements per angles file record.

An angles file data element is a 12-byte structure containing three angles that define the satellite and solar viewing geometry for a given pixel. Thus the length of an angles file record in bytes is given by:

ANG_REC_LEN = 12 * ANG_ELEM_PER_LINE

The 12 bytes consist of three 32-bit floating-point variables: SATZEN, SOLZEN, and AZIMUTH corresponding to the satellite zenith, the solar zenith, and the satellite/solar azimuth angles respectively (Figure B-1). Note: Angle files were generated on a VMS computer. To interpret these floating-point numbers on a UNIX machine it is necessary to convert from VMS to IEEE floating-point formats. Most UNIX operating systems provide a utility to perform this conversion. Angle measurement conventions are as follows:

SATZEN	Scene satellite zenith angle, 0^{o} - 90^{o}.
SOLZEN	Scene solar zenith angle, 0^{o} - 180^{o}.
AZIMUTH	Relative angle between the solar and satellite azimuth angles, 0^{o} - 359^{o}. When AZIMUTH = 0^{o}, the sun is directly behind the satellite (i.e., the viewed point, the satellite, and the sun are collinear). When AZIMUTH = 180^{o}, the satellite is looking directly into the sun (the satellite squints to compensate).

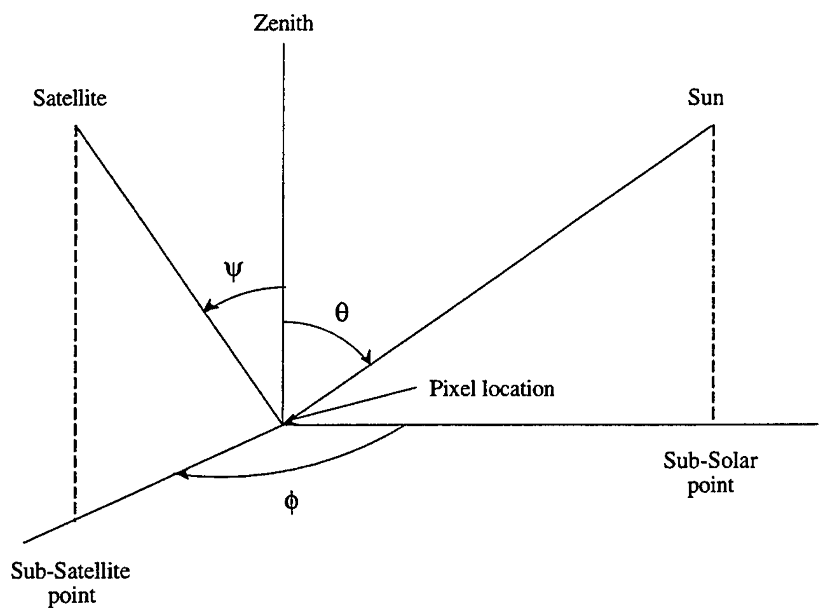


Figure B-1. Satellite-Earth-Solar Geometry (after Taylor and Stowe, 1984)

ψ - satellite zenith angle

θ - solar zenith angle

ϕ - sun-satellite azimuth angle

Level 2: **Nephanalysis Products**

Nephanalysis products are stored as bit-encoded byte values known as MCF (cloud Mask and Confidence Flag). MCF filenames as they appear on tape have the following naming convention:

SSS_MCF_ROI_DDD_HH.Dat or .Tif

where

SSS - Satellite identifier:
- F10 DMSP F-10
- F11 DMSP F-11
- N11 NOAA-11
- N12 NOAA-12
- G04 GMS-4 (Japan)

MCF - a constant that identifies the file as an MCF file
ROI - Region of Interest for which the product is valid:
EAS for East Asia Area
DDD - Julian day for which the product is valid
HH - UTC hour for which the product is valid
Dat - Raw product file format
Tif - TIFF file format

File and Record Structure

Level 2 processing is performed on square arrays of image pixels, therefore the size of the resultant MCF product files is an integral number of the analysis array size. MCF files contain fixed-length records, the number and size of which depends on both the size of the corresponding image files and the satellite type. The following table specifies how to determine the record size and number of records in an MCF file. Let NCOLS and NROWS be the number of columns and rows, respectively, in the corresponding satellite image file; then:

If the image satellite id is:	Then the MCF file record size is:	And the number of lines is:
DMSP F10 or F11	NCOLS - MOD(NCOLS, 16)	NROWS - MOD(NROWS, 16)
NOAA 11 or 12	NCOLS - MOD(NCOLS, 32)	NROWS - MOD(NROWS, 32)
GMS 4	See Associated RE File or TIFF Header	See Associated RE File or TIF Header

where MOD is the FORTRAN modulus function (e.g., if an F10 pass has 1465 columns per scan line, then the MCF record size is 1456). The MCF file is stored in Tagged Image File Format (TIFF), therefore an alternative way to determine file dimensions is to read the TIFF header and examine the width and height fields.

The format of an MCF file is the same regardless of the satellite platform it was derived from. The first byte of the first record of the MCF file corresponds to the first byte of the first record in the corresponding image data file. Across each scan line there is a one-to-one correspondence between the image and MCF files out to the number of bytes computed above for each record. As can be seen in the above table, the MCF and image file sizes are not always the same. However, the two files are always aligned with respect to the upper-left corner of each.

There is one 8-bit MCF byte per analyzed image pixel. MCF bytes are bit-packed according to the following convention:

Bit 0 (least significant) is the cloud/no-cloud bit. If bit 0 is off, the corresponding image pixel is clear; if bit 0 is on, it is completely cloudy.

Bit 1 is the low cloud bit. If bit 1 is on, the pixel contains low cloud as determined by an appropriate spectral (or other) signature test.

Bit 2 is the thin cirrus cloud bit. If bit 2 is on, the pixel contains cirrus as determined by an appropriate spectral (or other) signature test.

Bit 3 is the cumulonimbus bit. If bit 3 is on, the pixel contains thunderstorm clouds.

Bit 4 is the partly cloudy bit. If bit 4 is on, the pixel is partly cloudy. If bit 4 is on, bit 0 is off. DMSP data are used exclusively to determine partly cloud conditions.

Bit 5 is the bad data bit. It is set whenever satellite data are missing or unreliable. If set, all other bits should be ignored.

Bits 6 and 7 contain the confidence level attached to the accuracy of the cloud/no-cloud decision for the corresponding cloudy image pixel. Confidence levels are rated as 0 for missing data, 1 for low confidence, 2 for mid-level confidence, and 3 for high confidence.

Low cloud, thin cirrus, and cumulonimbus conditions are always associated with completely cloudy conditions (i.e., bit 0 will always be on in the presence of one or more of these conditions). Cloud level and cloud type are not detected under partly cloudy conditions (i.e., if bit 4 is on, bits 1 through 3 will be off).

Example:

MCF byte	1	1	0	0	0	1	0	1	(C5 in hex)
bit position	7	6	5	4	3	2	1	0	

The corresponding image pixel is classified as cloud covered (bit 0) with thin cirrus (bit 2) that has been detected with a high level of confidence (bits 6 and 7).

Level 3: Layered Product

The layered product filename as it appears on tape has the following naming convention:

SAT_LYR_ROI_DDD_HH.DAT

where:

SAT - Satellite identifier:

F10	DMSP F-10
F11	DMSP F-11
N11	NOAA-11
N12	NOAA-12
G04	GMS-4 (Japan)

LYR is a constant that denotes the file is a layered product

ROI - Region of Interest:

EAS for the DNA East Asia Area (EASA)

DDD - Julian day

HH - GMT hour

File Structure

The layered product file contains 86175 (225 rows x 383 columns) record structures, each 51 bytes in length.

Record Structure

Each record contains data values valid for one grid point within a 383 (rows) by 225 (columns) two-dimensional grid. The grid is superimposed on a hemispheric secant polar stereographic map projection. Grid resolution is based on a whole mesh grid spacing of 381 km at 60 degrees latitude and nested grids are defined in terms of the number of grid cells that fit within a whole mesh grid. The layered product grid is a 1/16th mesh grid (i.e., 16 by 16 grid cells per whole mesh box.)

Table B-2 summarizes the contents of each record. Figure B-2 contains the C data structure that was used to create the data file.

Table B-2: Layered Product Record Structure

Field	Description	Units	Range	Missing or bad value	Byte length
1	Absolute 16th-mesh row number (i)		1-1024		2
2	Absolute 16th-mesh column number (j)		1-1024		2
3	SDB IR entry number			0	2
4	Julian day (yyddd)				4
5	UTC (hhmm)		0-2359		2
6-9	Cloud temperature variance for each layer	GS*100			8
10-13	Cloud top temperature for each layer	GS*100			4
14-17	# pixels in each layer			0	4
18-21	Cloud type for each layer		0-1		4
22-25	# low cloud pixels in each layer				4
26-29	# thin cirrus pixels in each layer				4
30-33	# precipitating-cloud pixels in each layer				4
34	Sunrise time		0-235		1
35	Sunset time		0-235		1
36	Satellite platform ID				1
37	# pixels in grid box				1
38	# data dropouts in grid box				1
39	# partially cloud-filled pixels				1
40	Pad				1

```
/* Layering output structure

   Daniel Peduzzi (AER) 9/27/94
     structure content by Robert P. d'Entremont (AER) 9/1994
*/

#ifndef NCLASSES
#  define NCLASSES (4)
#endif

#ifndef _LAYER_OUTPUT
#define _LAYER_OUTPUT

#define BYTE unsigned char

typedef struct {

  short i;                          /* 16th-mesh absolute row (1-1024)              */
  short j;                          /* 16th-mesh absolute column (1-1024)           */

  short sdb_ir_entry;               /* SDB entry number corresponding to IR data    */
  int yyddd;                        /* Sensor data Julian day                       */
  short hhmm;                       /* Sensor data valid time (UTC) hhmm            */
  short layer_var[NCLASSES];        /* Temperature variance*100 for cloud layer i   */
  BYTE meantemp[NCLASSES];          /* Mean cloud top temperature for layer i       */
  BYTE n_layer_pix[NCLASSES];       /* Total # pixels in layer i                    */
  BYTE cloud_type[NCLASSES];        /* Cloud type for layer i (1 or 2)              */
  BYTE low_cloud[NCLASSES];         /* # low cloud pixels in layer i                */
  BYTE thin_cirrus[NCLASSES];       /* # thin cirrus pixels in layer i              */
  BYTE precip[NCLASSES];            /* # precipitating-cloud pixels in layer i      */

  BYTE sunrise;                     /* Sunrise time (UTC) (0-235)                   */
  BYTE sunset;                      /* Sunset time (UTC) (0-235)                    */
  BYTE vid;                         /* Satellite vehicle (platform) ID              */
  BYTE num_pixels;                  /* Total # of pixels in 16th-mesh box           */
  BYTE dropouts;                    /* Total # of data dropouts in 16th-mesh box    */
  BYTE partial;                     /* Total # of partially-cloud-filled pixels     */
  BYTE dummy;                       /* Pad                                          */

} LAYER_OUTPUT;

#undef BYTE

#endif
```

Figure B-2: Level 3 data structure

Level 4: Integrated Product

The integrated product filename as it appears on tape has the following naming convention:

ALL_IAN_ROI_DDD_HH.Dat

where

ALL and IAN are constants (Integrated ANalysis from ALL sensors)
ROI - Region of Interest for which the product is valid
Possible values:
EAS for the East Asia Area
DDD - Julian day for which the integrated product is valid
HH - GMT hour for which the integrated product is valid

File Structure

The integrated product file contains 86,175 records (225 columns by 383 rows), each 64 bytes in length.

Record Structure

Each record contains data values valid for one grid point within a 383 (rows) X 225 (columns) 2-D grid. The grid is superimposed on a hemispheric secant polar stereographic map projection. Grid resolution is based on a whole-mesh grid spacing of 381 km at 60^{o} latitude and nested grids are defined in terms of the number of grid cells that fit within a whole mesh grid. The integrated product grid is a $1/16^{th}$ mesh grid (i.e., 16 X 16 cells per whole mesh box).

Table B-3 summarizes the contents of each record. All values are 16-bit integers. Figure B-3 contains the C data structure used to create the output file.

Table B-3. Integrated Product Record Structure

Field	Description	Units	Range	Missing or bad value	Comments
1	Absolute 16th-mesh column number (i)		227 - 451		
2	Absolute 16th-mesh row number (j)		13 - 395		
3	Number of cloud layers in (i,j)		0 - 4	-999	
4	Total cloud fraction for (i,j)	Percent	0 - 100	-999	
5-8	Cloud fraction by layer for (i,j)	Percent	0 - 100	-999	
9-12	Cloud top temperature by layer	K*10	2000-3275	-999	
13-16	Cloud top height by layer	Meters	0-13500	-999	
17-20	Cloud type by layer		0 - 9	-999	See Table B-4
21	Total cloud fraction error for (i,j)	Percent	0 - 100	-999	
22-25	Layer cloud fraction error for (i,j)	Percent	0 - 100	-999	
26-29	Layer confidence flags for (i,j)	Flag*10	10 - 30	-999	Discrete values for low to high confidence
30-32	Database entry numbers for input satellite analyses				Corresponds to directory names on tar tape

Table B-4. Cloud Type Codes

Cloud Type Code	Cloud Type
0	No Cloud
1	Cirrus
2	Cirrostratus
3	Altocumulus
4	Altostratus
5	Stratocumulus
6	Stratus
7	Cumulus
8	Cumulonimbus
9	Nimbostratus

```
/* EASA definitions */

#define NLINE 383
#define NCOL 225
#define NLYRS 4


#define MIN_I 227
#define MIN_J 13
```

```
typedef unsigned char byte;

/* integration output structure */

typedef struct {
   short  i;                        /* absoulute 16th mesh coord        */
   short  j;
   short  nlayers;                  /* number of layers                 */
   short  fraction;                 /* total cloud fraction             */

   short  lyr_frc[NLYRS];           /* layer cloud fraction             */
   short  t_cld[NLYRS];             /* layer cloud top temp (K*10)      */
   short  z_cld[NLYRS];             /* layer cloud top height (m)       */
   short  cld_typ[NLYRS];           /* layer cloud type                 */

   short  error;                    /* total cloud amount error         */
   short  lyr_err[NLYRS];           /* layer cloud amount error         */
   short  conf[NLYRS];              /* layer confidence measure         */
   short  sdb_entry[3];             /* input entry number(s)            */
} INTEGRATION;
```

Figure B-3: Integration output data structure

Appendix C

Data Extraction Guide

****SERCAA DATA SET RELEASE TO DNA****

****
What should I have ?

DNA_RELEASE.TXT	This document.
(2) 8 mm D8-112 tapes	One tape, labeled DNA MAR93 IA, contains the SERCAA Integrated Analysis (SIA) data files. The other tape, labeled DNA MAR93 ENTRIES contains the Related Entry (RE) data (which consists of Satellite, Latitude/Longitude, Angles(Geometry) and Product(cloud mask) data files.

****
What type of tape drive was used ?

A SUN Exabyte EXB-8500 8 mm tape drive recording in high density mode (5 gig).

****
What utility was used to create the release tapes ?

The data were placed on the tapes using a SUN SPARC II running SUN OS 4.1.2. The following tar command syntax was used:

```
sun% tar cvBf /dev/nrst8 somedirectory
```

****
How are the data arranged on the release tape ?

The data on the SIA tape are contained in four tar files. Each of these tar files represents a directory that contains all the SIA data for a particular day (day 93081 through day 93089). Each directory name follows the convention:

CYYJJJ

where:

C = century (9 for 19XX)
YY = year
JJJ = Julian day

A SIA file and SIA SDB information file exists for each hour that an analysis was performed. Each SIA file has been named using the following convention:

Positions 1-4 Platform:

all_ = All satellite platforms are used to create a SIA.

Positions 5-8 Type of file:

ian_ = integrated analysis file
sdb_ = SERCAA data base (SDB)
information file

Positions 9-12 Region of interest:
(Given in 16th-mesh coordinates)
eas_ = East Asia Area (EASA). (i,j) = (227,13) to (451,395)
can_ = Canada Area (CANA). (409,597) to (557,711)
cns_ = Central, Northern South America Area (CNSA).
(413,877) to (651,1011)
emd_ = Eastern Mediterranean, Desert Area (EMDA).
(731,353) to (863,505)

Positions 13-16 Julian day:
081_ = Julian day 81 etc. ...

Positions 17-18 Hour:
00 = SIA for hour 00 etc. ...

Positions 19-22 Extension:
.dat = raw-format file extension

Example: all_ian_eas_081_10.dat

The RE tape contains the tar files. Each of these tar files represent a directory that contains all the related data used as input to create at least one of the SIA data files. Each directory name follows the convention:

ENTRY/

where:
ENTRY = the SDB entry number

Each RE file has been named following these guidelines:

Positions 1-4 Platform:
n11_ = NOAA N_11
n12_ = NOAA N_12
f10_ = DMSP F_10
f11_ = DMSP F_11
g04_ = GMS-4

Positions 5-8 Type of file:
001_ = satellite data channel 1
002_ = satellite data channel 2
...
...
005_ = satellite data channel 5
lat_ = latlon data
ang_ = angles data
mcf_ = cloud mask data
sdb_ = SDB information file

Positions 9-12 Area of data:
eas_ = East Asia Area (EASA)
can_ = Canada Area (CANA)
cns_ = Central and Northern South America Area (CNSA)

emd_ = Easter Mediterranean, Desert Area (EMDA)

Positions 13-16 Julian day:

081_ = Julian day 81 etc. ...

Positions 17-18 Hour:

00 = hour of the data

Positions 19-22 Extension:

.dat = raw data
.tif = tif formatted data

Examples:

f10_001_eas_150_14.tif
f10_002_eas_150_14.tif
f10_lat_eas_150_14.tif
f10_ang_eas_150_14.tif
f10_mcf_eas_150_14.tif
f10_sdb_eas_150_14.tif

Refer to separate listing sheet labeled MAR93.IA.TAR.LIST for a listing of the IA tape contents. Run the provided script, "list_tar", to generate a listing of the RE tape.

****

What are related data items ?
What is the SDB entry number ?
What are related entries ?

The SDB registration process is a process that automatically places descriptive data items about a satellite scan into the SDB. The SDB registration process allocates a group of unique entry numbers to be used as place holders for all of the related data items for a given satellite scan. The related data items consists of satellite, latitude/longitude, angles (Geometry) and product(cloud mask) data. As an example, if a DMSP F_11 scan was to be registered in the SDB, the registration process would request for a group of five contiguous entry numbers(i.e. 1001-1005). These five entry numbers would be used as place holder for the following related data items:

1001	f11 visible channel
1002	f11 infrared channel
1003	latitude/longitude data
1004	angles(geometry) data
1005	product data

The "SDB entry number" is the first entry number of the group of entry numbers provided by the registration process. The first entry number is used to "key" into the related data items for that group. In the example provided above the SDB entry number would be 1001.

This release process uses the SDB entry number in each group to logically divide the data into separate directories (i.e. the directory name is first SDB entry number for each group of entry numbers). Using the example provided above the directory named "1001/" contains all the related data items for that group (i.e. the directory contains the data for entry 1001 through entry 1005).

To build a SIA it is necessary to use as input, related data items from one or more satellite scans and/or satellite platforms. The SDB entry number is used to keep track of all inputs to the SIA. The list of related entries are given as SDB entry numbers.

****

How do I get a particular SIA data set ?

You must use the UNIX tar utility to extract the data from the tape. By using the tar utility you may extract individual files or the entire directory. Use the MAR93.IA.TAR.LIST to determine where to position the tape and then use the appropriate tar command to extract the files you want. For example, if you want to extract all of the SIA data files from the first and second tar files, the following commands could be used:

```
% pwd
/users/smith
% mkdir data
% cd data
% tar xvf /dev/rst8 993147
% tar xvf /dev/rst8 993148
```

Upon completion all of the SIA data for day 81 would reside in directory /users/smith/data/993081 and all the SIA data for day 81 would reside in directory
/users/smith/data/993081.

****

What is the SDB information file ?

The SDB information file is a text file containing selected SDB record items that help describe the actual data. The SIA SDB information file shows what data went into creating the SIA by listing the related entries. The RE SDB information file lists information about the satellite images, the latlon file, the angles file and the product file(s).

The following is an example SIA SDB information file:

```
[IA]
ZULU_YYJJJ:=93081                          : Year, Julian day of SIA
ZULU_HH:=10                                : Hour of SIA
ROI:=EAS                                   : Region of Interest
NUM_RELATED_LAYER:=3                       : Number for related entries
RELATED_LAYER_1:= 4148                     : 1st related SDB entry number
RELATED_LAYER_2:= 7199                     : 2nd related SDB entry number
RELATED_LAYER_3:= 8988                     : 3d related SDB entry number
TDISK:=SDB_Int:
TDIR:=[SERCAA.DATA.993081]
FILE_IA_1:=ALL_IAN_EAS_081_10.Dat          : SIA file name
SDB_SET:=MAR93                             : Set identifier March of 1993
```

The following is an example RE SDB information file:

```
[SATIMG]
SAT_CODE:=16                               : Satellite code
ZULU_YYJJJ:=93081                          : Year, Julian day of scan
ZULU_HHMMSS:=82252                         : Time of scan
NUM_LINES:=1375                            : Number of lines
```

```
ELEM_PER_LINE:=409                                              : Elements per line
BYTES_PER_ELEM:=1                                               : Bytes per element
7199:=AVH$005:[SERCAA.DATA.993081]N11_001_EAS_081_08.TIF        : Channel 1 file
7200:=AVH$005:[SERCAA.DATA.993081]N11_002_EAS_081_08.TIF        : Channel 2 file
7201:=AVH$005:[SERCAA.DATA.993081]N11_003_EAS_081_08.TIF        : Channel 3 file
7202:=AVH$005:[SERCAA.DATA.993081]N11_004_EAS_081_08.TIF        : Channel 4 file
7203:=AVH$005:[SERCAA.DATA.993081]N11_005_EAS_081_08.TIF        : Channel 5 file

[LATLON]
LL_REC_LEN:=204                                                 : Record length in bytes
LL_LINE_INTERVAL:=1                                             : Sub-sample line interval
LL_ELEM_INTERVAL:=8                                             : Sub-sample element interval
LL_ELEM_PER_LINE:=51                                            : Latlon pairs per line
LL_FILE:=AVH$005:[SERCAA.DATA.993081]N11_LAT_EAS_081_08.DAT     : latitude/longitude file

[ANGLES]
ANG_REC_LEN:=612                                                : Record length in bytes
ANG_LINE_INTERVAL:=1                                            : Sub-sample line interval
ANG_ELEM_INTERVAL:=8                                            : Sub-sample element interval
ANG_ELEM_PER_LINE:=51                                           : Angles triplets per line
ANG_FILE:=AVH$005:[SERCAA.DATA.993081]N11_ANG_EAS_081_08.DAT    : Angles file

[PRODUCT]
7206001:=sdb$prd:[SERCAA.DATA.993081]N11_MCF_SET_081_08.TIF     : Cloud mask file
```

****

How do I know which RE data went into a particular SIA ?

There are two ways to determine which RE data sets went into a particular SIA. The first way is reference the SIA SDB information file. Each "RELATED_LAYER" listed is a reference, by SDB entry number, to the RE data. Use the referred SDB entry number to retrieve the related data from the RE data tape.

For example, refer to the above SIA SDB information file. The "RELATED_LAYERED_1:=4148" line implies that SDB entry number 4148 and the related data items for entry 4148 (along with SDB entry numbers 7199 and 8988) were used to create "ALL_IAN_EAS_081_10.Dat".

The second way is to read the header information from the SIA file (Please refer to the DATA_DESCRIPTION).

****

How do I get the RE data files ?

Once you have examined the SIA SDB information file and you have identified the related entry numbers, you must use the UNIX tar utility to extract the data from the tape. By using the tar utility you may extract individual files or the entire directory. Use a tape contents list generated using the "list_tar" script to determine where to position the tape and then use the appropriate tar command to extract the files you want. For example, if you want to extract all of the RE data files from the first tar file, the following commands could be used:

```
% pwd
/users/smith
% mkdir data
% cd data
% tar xvf /dev/rst8 7199
```

Upon completion of this command all of the RE data related to SDB entry number 7199 would reside in directory /users/smith/data/7199.

****

For the following question please refer to the example SDB information files as needed.

****

What is the format of the satellite data and how do I access it?

The dimensions of the satellite data are defined by the three parameters, NUM_OF_LINES, ELEM_PER_LINE and BYTES_PER_ELEM . To access the data use the following logic.

If the file extension is ".dat"
then use the appropriate C or FORTRAN read statements.

If the file extension is ".tif"
then use a tiff reader or tiff library (you may view the
images by using the public domain application, XV).

For a detailed explanation, refer to Appendix B.

****

What is the format of the latlon data and how do I access it?

The latlon data are sub-sampled. The dimensions are defined LL_LINE_INTERVAL, LL_ELEM_INTERVAL and LL_ELEM_PER_LINE. LL_ELEM_PER_LINE defines the number of longitude/latitude pairs per line. Each pair is four bytes (two bytes lon, two bytes lat). To access the data use the appropriate C or FORTRAN read statements.

For a detailed explanation, refer to Appendix B.

****

What is the format of the angles data and how do I access it?

The angles data are sub-sampled. The dimensions are defined by ANG_LINE_INTERVAL, ANG_ELEM_INTERVAL and ANG_ELEM_PER_LINE. ANG_ELEM_PER_LINE
defines the number of triplets (satellite-zenith/solar-zenith/azimuth) per line. Each item in the triplet is a float data type. To access the data use the appropriate C or FORTRAN read statements.

For a detailed explanation, refer to Appendix B.

P601

Data Save Documentation Report No. 3

ADVANCED GEOPHYSICAL ENVIRONMENT SIMULATION TECHNIQUES

Task 1: Satellite Data Sets for Worldwide Cloud Prediction

This data documentation report covers data set generation
for the DNA region of interest:

East Asia Area (EASA)

for the period:

22-31 July 1993

Contract Number F19628-94-C-0106

issued by:

Electronic Systems Division
Air Force Systems Command
Hanscom AFB, MA 01731

Submitted by:

Atmospheric and Environmental Research, Inc.
840 Memorial Drive
Cambridge, MA 02139

22 December 1994

David B. Hogan
Gary B. Gustafson
Principal Investigators

1.0 Introduction

This Data Documentation Report provides a description of the third data save made in accordance with the revised statement of work for Satellite Data Sets for Worldwide Cloud Prediction Models. It is intended to provide a description of the data set, its format, how it was gathered and processed, and a description of the algorithms used to generate it. The data set consists of raw satellite data and analyzed products produced by the SERCAA cloud analysis algorithms. The period covered is 22-31 July 1993 for the DNA region of interest: East Asia Area (EASA). This region covers the following (i,j) 16^{th} mesh grid coordinates: 227,13 - 451,395. All available data from those dates are included. These data were processed specifically for DNA using software developed from the SERCAA cloud analysis algorithms described by Gustafson et. al (1994). Substantial modifications were required to the Cloud Layering and Analysis Integration modules to accommodate the high volume of data included in this data set. Two tapes are provided, one with Level 1, 2 and 3 products and the second with Level 4. Data formats for the Level 3 products differ slightly from those used in the March 1993 data set provided earlier (see Data Save Documentation Report No. 2, dated 30 November 1994).

2.0 Processing Environment

Satellite data processing for this data set used the SERCAA cloud analysis algorithms described by Gustafson et al. (1994). Multisource data from the DMSP F10 and F11, NOAA-11 and NOAA-12, and GMS-4 satellites were used. Data sources were as follows: DMSP - National Geophysical Data Center (NGDC), Boulder, CO; NOAA - National Climatic Data Center (NCDC), Ashville, NC; GMS - Sea Space Corp., San Diego, CA. All data were obtained by the Phillips Laboratory and were received on tape in various formats. All data processing was performed on the Air Force Interactive Meteorological System (AIMS) at the Phillips Laboratory. The SERCAA cloud analysis algorithms use four levels of data processing as summarized in Figure 1.

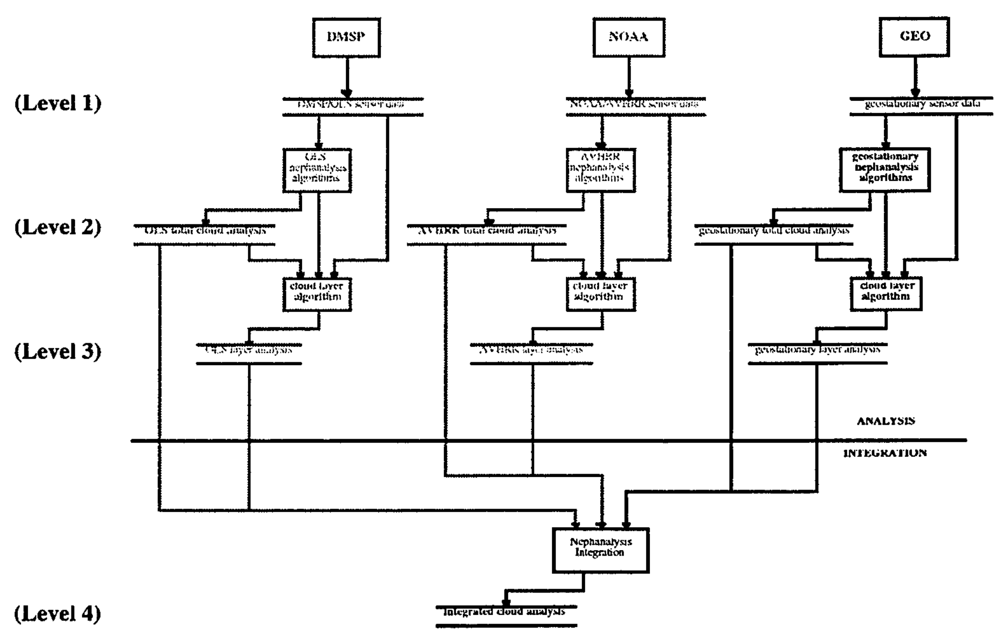


Figure 1 SERCAA data flow and processing levels

Level 1 processing consists of data ingest. Tape data are processed through separate ingest programs depending on the data source and format. All data are then stored in a standard format in the original satellite scan projection. The format consists of flat files where the number of elements correspond to the number of pixels in the satellite scan line and the number of rows corresponds to the number of scan lines. Data are maintained on AIMS through the SERCAA Database (SDB) management software. Level 1 data products consist of separate files for each sensor channel plus two additional files containing Earth location and satellite/solar geometry information. Satellite data characteristics are summarized in Table 1. In cases where visible and infrared channel resolution differ, the higher resolution data are subsampled to match the coarser resolution data (e.g., GMS visible data are subsampled by a factor of four to match the IR data resolution). Earth location data consist of latitude-longitude pairs that are maintained at a subsampled resolution relative to the satellite data. For each sensor scan line, one latitude-longitude pair is provided for every n^{th} pixel, where n varies with satellite. Geometry information are also subsampled in the same ratio as the Earth location information and consist of three angles: satellite zenith, solar zenith, and sun-satellite azimuth. Ingest products are described more completely in Section 2 of Gustafson et al. (1994).

Table 1. Sensor Channel Data Attributes During SERCAA

Satellite	Sensor	Channel (μm)	Data Format	Resolution[1] (km)	Bits per Pixel[2]	Pixels per Scan Line
DMSP	OLS	0.40-1.10	counts	2.7	6	1464
		10.5-12.6	EBBT	2.7	8	1464
NOAA	AVHRR	0.58-0.68	percent albedo	4.0	10	409
		0.72-1.10	percent albedo	4.0	10	409
		3.55-3.93	EBBT	4.0	10	409
		10.3-11.3	EBBT	4.0	10	409
		11.5-12.5	EBBT	4.0	10	409
GMS	VISSR	0.5-0.75	counts	1.25	6	10000
		10.5-12.5	EBBT	5.0	8	2500

[1]Sensor resolution at satellite subpoint that will provide global coverage.

[2]AVHRR radiance data are transmitted at 10-bit resolution, however, the SERCAA development system could only accommodate 8-bit brightness temperature data (although the full 10-bit resolution is used in the radiance-to-brightness-temperature transformation).

Level 2 processing consists of sensor-specific nephanalysis algorithms. Level 1 sensor data from DMSP, NOAA, and the GMS geostationary satellites are processed through separate algorithms as indicated in Figure 1. Each time data from a new satellite pass are ingested, they are analyzed through the appropriate nephanalysis algorithm and results are placed in a Level 2 output file. One output file is generated for each nephanalysis run and nephanalysis results are stored in the original satellite scan projection with one byte of information for each pixel. Each byte is bit-packed according to the map in Table 2. For each set of Level 1 products generated from a satellite pass, one Level 2 product file is generated.

Table 2. Cloud Analysis Algorithm MCF File Bit Assignments

Bit	Assignment	Description
0	Cloud Mask	ON = Cloud-Filled OFF = Cloud-Free
1	Low Cloud	ON = Low Cloud Found
2	Thin Cirrus Cloud	ON = Thin Cirrus Cloud Found
3	Precipitating Cloud	ON = Precipitating Cloud Found
4	Partial Cloud	Only used by DMSP algorithm
5	Data Dropout	ON = Missing or Unreliable Data
6	Confidence	0 = Missing Data; 1 = Low;
7	Flag	2 = Middle; 3 = High

Level 3 processing uses Level 1 and 2 products as input to segment the cloudy regions into vertical cloud layers and to classify different cloud types. It also remaps the data from the individual satellite projections to the AFGWC standard polar stereographic map projection (Hoke et al., 1981) at 16th mesh grid resolution. The EASA region of interest processed for the July 1993 data set have the following (i,j) 16th mesh grid coordinates: $395 \leq i \leq 451$, $13 \leq j \leq 227$. Level 3 products are generated for each 16th mesh grid cell and contain the information in Table 3. A maximum of four cloud layers can be identified for each grid cell. One Level 3 file is created for each set of Level 1 and 2 products. All Level 1, 2, and 3 products associated with a single satellite pass are related through SDB and are provided on the DNA tapes as a set. Note that for the EASA region, all Level 3 files are a fixed size of 225x383 grid cells.

Table 3. Cloud Typing and Layering Output

Parameter	Description
i	16th mesh i coordinate for Grid Cell
j	16th mesh j coordinate for Grid Cell
sdb_ir_entry	SDB entry number of input IR sensor data
ddd	Sensor data Julian date
hhmm	Sensor data valid time (UTC)
layer_var(4)	Cloud top IR variance of pixels in each layer
num_pixels	Total number of satellite pixels in 16th mesh grid cell
n_layer_pix(4)	Total number of pixels in each layer
meantemp(4)	Cloud top mean IR Temperature of pixels in each layer
cloud_type(4)	Cloud type of each layer
low_cloud(4)	Number of low cloud pixels in this layer detected by cloud analysis algorithm
thin_cirrus(4)	Number of thin cirrus pixels in this layer detected by cloud analysis algorithm
precip(4)	Number of precipitating cloud pixels in this layer detected by cloud analysis algorithm
sunrise	Local sunrise time (UTC)
sunset	Local sunset time (UTC)
vid	Satellite vehicle (platform) ID
dropouts	Number of bad data pixels in 16th mesh grid cell
partial	Number of partial cloud pixels detected by DMSP cloud analysis algorithm

Level 4 processing is a clock driven process with one new Level 4 integrated analysis performed each hour. Thus, integration is differentiated from the Level 1, 2, and 3 products that are event-driven (i.e., resulting from the ingest of a new satellite pass). The integration module operates on the most recent Level 3 gridded products available from each satellite source (i.e., NOAA, DMSP, GMS). Like Level 3 products, the Level 4 output files conform to the AFGWC 16th mesh grid structure; output parameters for each grid cell are summarized in Table 4.

Table 4 Analysis Integration Processed Parameters

Parameter	Description
i	16th mesh i (column) coordinate
j	16th mesh j (row) coordinate
nlay	Number of Cloud Layers
cftot	Total Cloud Fraction
cf(4)	Layer Cloud Fraction
ctt(4)	Layer Cloud Top IR Temperature (K)
ctz(4)	Layer clout top height (m)
ity(4)	Layer Cloud Type
ecft	Estimated Error in Total Cloud Fraction
ecf(4)	Estimated Error in Layer Cloud Fraction
icf(4)	Analysis Confidence Flag Index For Each Layer
sdb(3)	SDB entry number of input analyses (NOAA, DMSP, GMS)

3.0 Tape Format

All data for the July 1993 EASA data save are contained on two 8 mm tapes written in UNIX tar format. The first tape, labeled: DNA JUL93 ENTRIES (RE), contains all the Level 1-3 products. The second tape, labeled: DNA JUL93 IA, contains all Level 4 products. The size of the combined Level 1, 2 and 3 products is approximately 2.7 Gbytes and the Level 4 products occupy 1.3 Gbytes. In addition to the two tapes, hard-copy listings of the contents of the Level 4 tape are also provided. The corresponding listing of the Level 1-3 tape is very large, so a UNIX script is provided to generate a listing at the user's site. It may be useful to place the listing file generated by the script into an edit program to scan and search it quickly. The listings are required to locate specific data sets on the tapes.

Level 1-3 products are generated for each new pass of satellite data received during the period of the data save. Appendix A contains a chronological list of each satellite pass used to produce the July 93 data sets. All available data for the period covered were included; any gaps in the data list are due to either missing or bad data. Numerous DMSP orbits contained periodic data dropouts as illustrated in Figure 2, the most severely affected files were removed from the data set. For data archiving purposes all Level 1-3 products associated with a given satellite pass were placed in a single directory and subsequently placed on tape as a single tar file. Thus the first tape contains a series of several hundred tar files; each file contains all Level 1-3 products associated with a single satellite pass. Level 4 files are grouped on the second tape by day, thus for the July data save there are ten tar files on the Level 4 tape that each contain all Level 4 output files for each of the ten days 93203-93212 (22-31 July 1993). For each set of

Level 1-3 products, and for each Level 4 file there is also an SDB Information File. These files contain descriptive metadata information extracted from the SERCAA Database that describe the relevant attributes of the SERCAA product files. For example, information files list the number of pixels in a scan line of satellite data and the number of scan lines in the file. Information on subsampling ratios for the Earth location and angles files are also contained there.

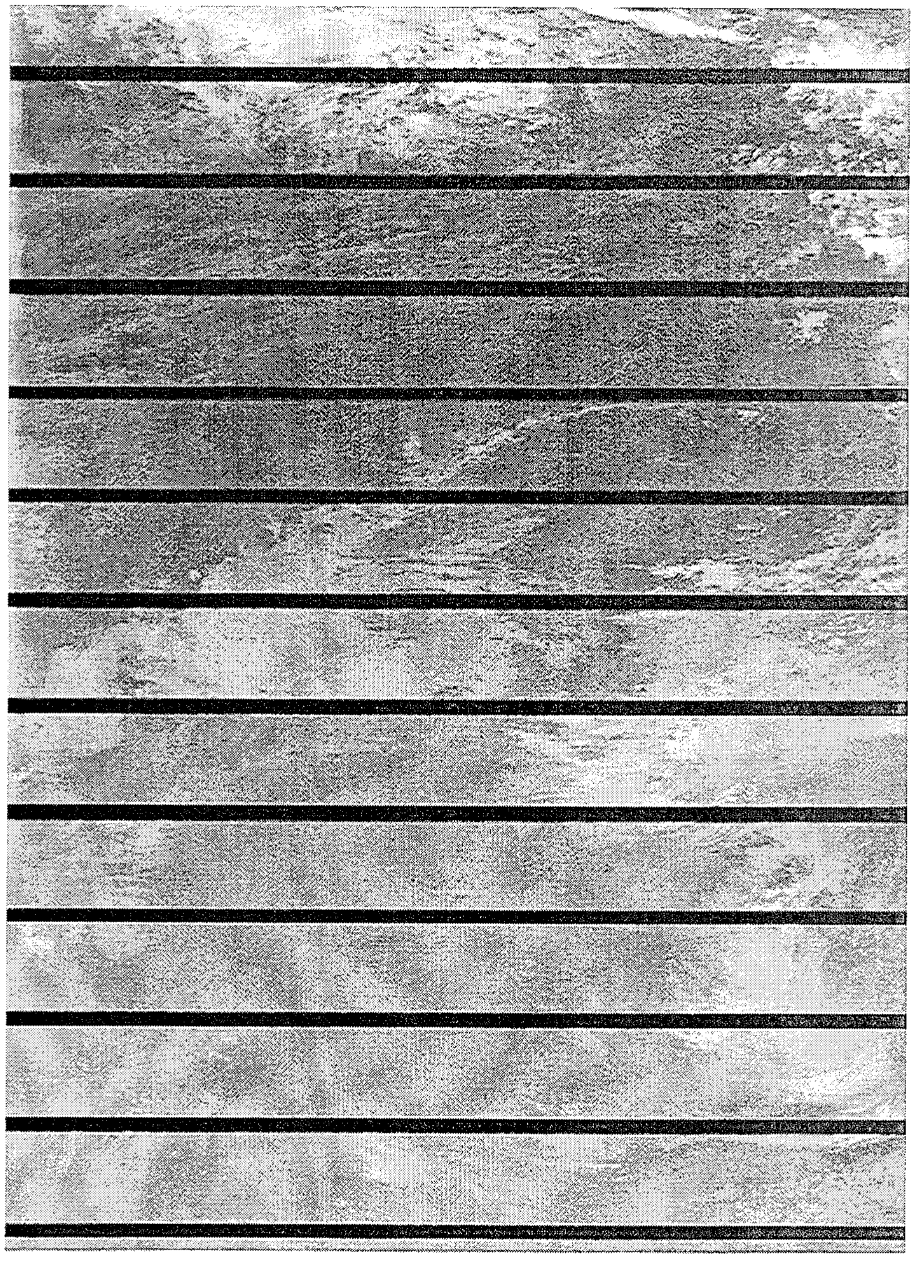


Figure 2 Sample of DMSP periodic data dropouts.

Detailed descriptions of the file formats used for each output level, and the associated information files, provided for the July 1993 save (Level 1, 2, 3, and 4) are provided in Appendix B. Appendix C provides a guide for extracting data sets from tape.

4.0 References

Gustafson, G.B., R.G. Isaacs, R.P. d'Entremont, J.M. Sparrow, T.M. Hamill, C. Grassotti, D.W. Johnson, C.P. Sarkisian, D.C. Peduzzi, B.T. Pearson, V.D. Jakabhazy, J.S. Belfiore, and A.S. Lisa, 1994: Support of Environmental Requirements for Cloud Analysis and Archive (SERCAA): algorithm descriptions. PL-TR-94-2114, Phillips Laboratory, Hanscom AFB, MA, ADA283240.

Hoke, J.E., J.L. Hayes, L.G. Renninger, 1981: Map projections and grid systems for meteorological applications. AFGWC-TN-79-003, Air Weather Service, Scott, AFB, IL.

Appendix A

Chronological List of Input Satellite Data

ENTRY	SATELLITE	DATE	TIME	ROI	ELES	LINES	RESLN
1176	NOAA_12	93203	002042	EAS	409	928	4.00
766	DMSP_f10	93203	011500	EAS	1465	2325	2.70
1088	NOAA_11	93203	070311	EAS	409	1800	4.00
1986	DMSP_f11	93203	075900	EAS	1465	1602	2.70
1200	NOAA_12	93203	081932	EAS	409	781	4.00
1991	DMSP_f11	93203	093500	EAS	1465	2177	2.70
1192	NOAA_12	93203	095621	EAS	409	1828	4.00
1184	NOAA_12	93203	113530	EAS	409	1092	4.00
771	DMSP_f10	93203	115100	EAS	1465	1365	2.70
1	GMS_4	93203	153400	EAS	1376	1010	5.00
6	GMS_4	93203	162700	EAS	1376	1010	5.00
11	GMS_4	93203	173400	EAS	1376	1010	5.00
1096	NOAA_11	93203	175339	EAS	409	918	4.00
16	GMS_4	93203	183400	EAS	1376	1010	5.00
1996	DMSP_f11	93203	192600	EAS	1465	1062	2.70
1104	NOAA_11	93203	193121	EAS	409	1992	4.00
21	GMS_4	93203	193400	EAS	1376	1010	5.00
26	GMS_4	93203	203400	EAS	1376	1010	5.00
31	GMS_4	93203	213500	EAS	1376	1010	5.00
1208	NOAA_12	93203	221848	EAS	409	1801	4.00
36	GMS_4	93203	222700	EAS	1376	1010	5.00
781	DMSP_f10	93203	230800	EAS	1465	747	2.70
41	GMS_4	93203	233300	EAS	1376	1010	5.00
46	GMS_4	93204	003500	EAS	1376	1010	5.00
51	GMS_4	93204	013500	EAS	1376	1010	5.00
56	GMS_4	93204	023500	EAS	1376	1010	5.00
61	GMS_4	93204	033300	EAS	1376	1010	5.00
66	GMS_4	93204	042800	EAS	1376	1010	5.00
1128	NOAA_11	93204	051435	EAS	409	1231	4.00
71	GMS_4	93204	053400	EAS	1376	1010	5.00
86	GMS_4	93204	063500	EAS	1376	1010	5.00
1120	NOAA_11	93204	065105	EAS	409	1872	4.00
91	GMS_4	93204	073500	EAS	1376	1010	5.00
2021	DMSP_f11	93204	082300	EAS	1465	2231	2.70
1112	NOAA_11	93204	083248	EAS	409	521	4.00
96	GMS_4	93204	083500	EAS	1376	1010	5.00
101	GMS_4	93204	093500	EAS	1376	1010	5.00
1232	NOAA_12	93204	093551	EAS	409	1722	4.00
2026	DMSP_f11	93204	100400	EAS	1465	710	2.70
1504	DMSP_f10	93204	102000	EAS	1465	1133	2.70
106	GMS_4	93204	102700	EAS	1376	1010	5.00
1224	NOAA_12	93204	111345	EAS	409	1412	4.00
76	GMS_4	93204	113400	EAS	1376	1010	5.00
116	GMS_4	93204	123500	EAS	1376	1010	5.00
1524	DMSP_f10	93204	125600	EAS	1465	2152	2.70
791	GMS_4	93204	133500	EAS	1376	1010	5.00
796	GMS_4	93204	143400	EAS	1376	1010	5.00
1509	DMSP_f10	93204	143500	EAS	1465	1109	2.70
801	GMS_4	93204	153400	EAS	1376	1010	5.00
806	GMS_4	93204	162700	EAS	1376	1010	5.00

ENTRY	SATELLITE	DATE	TIME	ROI	ELES	LINES	RESLN
811	GMS_4	93204	173400	EAS	1376	1010	5.00
1240	NOAA_11	93204	174207	EAS	409	730	4.00
816	GMS_4	93204	183400	EAS	1376	1010	5.00
2011	DMSP_f11	93204	191300	EAS	1465	827	2.70
1248	NOAA_11	93204	191921	EAS	409	1949	4.00
821	GMS_4	93204	193400	EAS	1376	1010	5.00
826	GMS_4	93204	203400	EAS	1376	1010	5.00
831	GMS_4	93204	213500	EAS	1376	1010	5.00
1368	NOAA_12	93204	215757	EAS	409	1491	4.00
836	GMS_4	93204	222700	EAS	1376	1010	5.00
876	GMS_4	93204	233300	EAS	1376	1010	5.00
1216	NOAA_12	93204	235729	EAS	409	2117	4.00
1514	DMSP_f10	93205	001300	EAS	1465	2159	2.70
841	GMS_4	93205	003500	EAS	1376	1010	5.00
846	GMS_4	93205	013500	EAS	1376	1010	5.00
851	GMS_4	93205	023500	EAS	1376	1010	5.00
856	GMS_4	93205	033400	EAS	1376	1010	5.00
861	GMS_4	93205	042800	EAS	1376	1010	5.00
1264	NOAA_11	93205	050240	EAS	409	1070	4.00
866	GMS_4	93205	053400	EAS	1376	1010	5.00
871	GMS_4	93205	063500	EAS	1376	1010	5.00
1272	NOAA_11	93205	063935	EAS	409	1854	4.00
1256	NOAA_11	93205	082026	EAS	409	781	4.00
136	GMS_4	93205	083500	EAS	1376	1010	5.00
1400	NOAA_12	93205	091510	EAS	409	1513	4.00
141	GMS_4	93205	093500	EAS	1376	1010	5.00
146	GMS_4	93205	102700	EAS	1376	1010	5.00
1384	NOAA_12	93205	105159	EAS	409	1660	4.00
151	GMS_4	93205	113400	EAS	1376	1010	5.00
156	GMS_4	93205	123500	EAS	1376	1010	5.00
161	GMS_4	93205	133500	EAS	1376	1010	5.00
166	GMS_4	93205	143400	EAS	1376	1010	5.00
171	GMS_4	93205	153400	EAS	1376	1010	5.00
176	GMS_4	93205	162700	EAS	1376	1010	5.00
1280	NOAA_11	93205	173040	EAS	409	545	4.00
181	GMS_4	93205	173400	EAS	1376	1010	5.00
186	GMS_4	93205	183400	EAS	1376	1010	5.00
1288	NOAA_11	93205	190723	EAS	409	1902	4.00
191	GMS_4	93205	193500	EAS	1376	1010	5.00
196	GMS_4	93205	203400	EAS	1376	1010	5.00
201	GMS_4	93205	213500	EAS	1376	1010	5.00
1424	NOAA_12	93205	213716	EAS	409	1121	4.00
206	GMS_4	93205	222700	EAS	1376	1010	5.00
1432	NOAA_12	93205	231450	EAS	409	2010	4.00
211	GMS_4	93205	233300	EAS	1376	1010	5.00
216	GMS_4	93206	003500	EAS	1376	1010	5.00
2046	DMSP_f10	93206	004300	EAS	1465	1639	2.70
1769	DMSP_f10	93206	012000	EAS	1465	2434	2.70
221	GMS_4	93206	013500	EAS	1376	1010	5.00
226	GMS_4	93206	023500	EAS	1376	1010	5.00
231	GMS_4	93206	033300	EAS	1376	1010	5.00
236	GMS_4	93206	042800	EAS	1376	1010	5.00

ENTRY	SATELLITE	DATE	TIME	ROI	ELES	LINES	RESLN
1320	NOAA_11	93206	045041	EAS	409	870	4.00
241	GMS_4	93206	053400	EAS	1376	1010	5.00
1312	NOAA_11	93206	062803	EAS	409	1831	4.00
246	GMS_4	93206	063500	EAS	1376	1010	5.00
251	GMS_4	93206	073500	EAS	1376	1010	5.00
1304	NOAA_11	93206	080805	EAS	409	1012	4.00
256	GMS_4	93206	083500	EAS	1376	1010	5.00
1464	NOAA_12	93206	085419	EAS	409	1291	4.00
261	GMS_4	93206	093500	EAS	1376	1010	5.00
266	GMS_4	93206	102700	EAS	1376	1010	5.00
1456	NOAA_12	93206	103013	EAS	409	1852	4.00
271	GMS_4	93206	113400	EAS	1376	1010	5.00
1774	DMSP_f10	93206	115600	EAS	1465	1653	2.70
276	GMS_4	93206	123500	EAS	1376	1010	5.00
1779	DMSP_f10	93206	133100	EAS	1465	2059	2.70
281	GMS_4	93206	133500	EAS	1376	1010	5.00
286	GMS_4	93206	143400	EAS	1376	1010	5.00
291	GMS_4	93206	153400	EAS	1376	1010	5.00
296	GMS_4	93206	162700	EAS	1376	1010	5.00
301	GMS_4	93206	173400	EAS	1376	1010	5.00
306	GMS_4	93206	183400	EAS	1376	1010	5.00
1328	NOAA_11	93206	185528	EAS	409	1850	4.00
311	GMS_4	93206	193400	EAS	1376	1010	5.00
316	GMS_4	93206	203400	EAS	1376	1010	5.00
1488	NOAA_12	93206	211649	EAS	409	789	4.00
321	GMS_4	93206	213500	EAS	1376	1010	5.00
326	GMS_4	93206	222700	EAS	1376	1010	5.00
1496	NOAA_12	93206	225337	EAS	409	1943	4.00
331	GMS_4	93206	233300	EAS	1376	1010	5.00
336	GMS_4	93207	003500	EAS	1376	1010	5.00
341	GMS_4	93207	013500	EAS	1376	1010	5.00
1784	DMSP_f10	93207	014900	EAS	1465	2329	2.70
346	GMS_4	93207	023500	EAS	1376	1010	5.00
356	GMS_4	93207	033300	EAS	1376	1010	5.00
361	GMS_4	93207	042800	EAS	1376	1010	5.00
1360	NOAA_11	93207	043841	EAS	409	607	4.00
1352	NOAA_11	93207	061627	EAS	409	1807	4.00
1344	NOAA_11	93207	075543	EAS	409	1212	4.00
1633	NOAA_12	93207	083320	EAS	409	1014	4.00
1296	DMSP_f11	93207	084600	EAS	1465	2143	2.70
1336	DMSP_f11	93207	092500	EAS	1465	1548	2.70
1641	NOAA_12	93207	100941	EAS	409	1855	4.00
1617	NOAA_12	93207	114945	EAS	409	835	4.00
1376	NOAA_11	93207	184335	EAS	409	1790	4.00
1392	NOAA_11	93207	202245	EAS	409	2129	4.00
1665	NOAA_12	93207	223231	EAS	409	1865	4.00
1673	NOAA_12	93208	001130	EAS	409	1906	4.00
1789	DMSP_f10	93208	025700	EAS	1465	2532	2.70
366	GMS_4	93208	053400	EAS	1376	1010	5.00
1416	NOAA_11	93208	060447	EAS	409	1751	4.00
371	GMS_4	93208	063500	EAS	1376	1010	5.00
376	GMS_4	93208	073500	EAS	1376	1010	5.00

ENTRY	SATELLITE	DATE	TIME	ROI	ELES	LINES	RESLN
1408	NOAA_11	93208	074322	EAS	409	1389	4.00
1681	NOAA_12	93208	081212	EAS	409	621	4.00
1713	DMSP_f11	93208	083300	EAS	1465	2017	2.70
381	GMS_4	93208	083500	EAS	1376	1010	5.00
1842	DMSP_f11	93208	091200	EAS	1465	1754	2.70
386	GMS_4	93208	093500	EAS	1376	1010	5.00
1689	NOAA_12	93208	094917	EAS	409	1811	4.00
391	GMS_4	93208	102800	EAS	1376	1010	5.00
396	GMS_4	93208	113400	EAS	1376	1010	5.00
1903	DMSP_f10	93208	123100	EAS	1465	2062	2.70
401	GMS_4	93208	123500	EAS	1376	1010	5.00
1908	DMSP_f10	93208	130900	EAS	1465	1589	2.70
406	GMS_4	93208	133500	EAS	1376	1010	5.00
411	GMS_4	93208	143400	EAS	1376	1010	5.00
416	GMS_4	93208	153400	EAS	1376	1010	5.00
421	GMS_4	93208	162700	EAS	1376	1010	5.00
426	GMS_4	93208	173400	EAS	1376	1010	5.00
1440	NOAA_11	93208	183145	EAS	409	1593	4.00
431	GMS_4	93208	183400	EAS	1376	1010	5.00
436	GMS_4	93208	193400	EAS	1376	1010	5.00
1448	NOAA_11	93208	201038	EAS	409	2102	4.00
441	GMS_4	93208	203400	EAS	1376	1010	5.00
446	GMS_4	93208	213500	EAS	1376	1010	5.00
1697	NOAA_12	93208	221133	EAS	409	1769	4.00
451	GMS_4	93208	222700	EAS	1376	1010	5.00
456	GMS_4	93208	233300	EAS	1376	1010	5.00
1705	NOAA_12	93208	235005	EAS	409	2098	4.00
461	GMS_4	93209	003500	EAS	1376	1010	5.00
466	GMS_4	93209	013500	EAS	1376	1010	5.00
471	GMS_4	93209	023500	EAS	1376	1010	5.00
476	GMS_4	93209	033500	EAS	1376	1010	5.00
481	GMS_4	93209	042700	EAS	1376	1010	5.00
486	GMS_4	93209	053400	EAS	1376	1010	5.00
1480	NOAA_11	93209	055304	EAS	409	1634	4.00
491	GMS_4	93209	063500	EAS	1376	1010	5.00
1472	NOAA_11	93209	073100	EAS	409	1541	4.00
496	GMS_4	93209	073500	EAS	1376	1010	5.00
501	GMS_4	93209	083500	EAS	1376	1010	5.00
1721	NOAA_12	93209	092842	EAS	409	1652	4.00
506	GMS_4	93209	093500	EAS	1376	1010	5.00
1898	DMSP_f11	93209	095800	EAS	1465	1931	2.70
511	GMS_4	93209	102700	EAS	1376	1010	5.00
516	GMS_4	93209	113400	EAS	1376	1010	5.00
521	GMS_4	93209	123500	EAS	1376	1010	5.00
526	GMS_4	93209	133500	EAS	1376	1010	5.00
1918	DMSP_f10	93209	133700	EAS	1465	2001	2.70
531	GMS_4	93209	143400	EAS	1376	1010	5.00
536	GMS_4	93209	153400	EAS	1376	1010	5.00
541	GMS_4	93209	162700	EAS	1376	1010	5.00
546	GMS_4	93209	173400	EAS	1376	1010	5.00
1529	NOAA_11	93209	181959	EAS	409	1357	4.00
551	GMS_4	93209	183400	EAS	1376	1010	5.00

ENTRY	SATELLITE	DATE	TIME	ROI	ELES	LINES	RESLN
556	GMS_4	93209	193400	EAS	1376	1010	5.00
1537	NOAA_11	93209	195832	EAS	409	2070	4.00
561	GMS_4	93209	203400	EAS	1376	1010	5.00
1913	DMSP_f11	93209	204700	EAS	1465	1487	2.70
1928	DMSP_f11	93209	212600	EAS	1465	2440	2.70
566	GMS_4	93209	213500	EAS	1376	1010	5.00
1729	NOAA_12	93209	215045	EAS	409	1358	4.00
571	GMS_4	93209	222700	EAS	1376	1010	5.00
1923	DMSP_f10	93209	231800	EAS	1465	1140	2.70
1737	NOAA_12	93209	232844	EAS	409	2052	4.00
576	GMS_4	93209	233300	EAS	1376	1010	5.00
581	GMS_4	93210	003500	EAS	1376	1010	5.00
586	GMS_4	93210	013500	EAS	1376	1010	5.00
591	GMS_4	93210	023500	EAS	1376	1010	5.00
596	GMS_4	93210	033300	EAS	1376	1010	5.00
601	GMS_4	93210	042800	EAS	1376	1010	5.00
606	GMS_4	93210	053400	EAS	1376	1010	5.00
1553	NOAA_11	93210	054118	EAS	409	1519	4.00
611	GMS_4	93210	063500	EAS	1376	1010	5.00
1933	DMSP_f11	93210	070800	EAS	1465	1719	2.70
1545	NOAA_11	93210	071838	EAS	409	1674	4.00
616	GMS_4	93210	073500	EAS	1376	1010	5.00
621	GMS_4	93210	083500	EAS	1376	1010	5.00
1802	NOAA_12	93210	090758	EAS	409	1442	4.00
626	GMS_4	93210	093500	EAS	1376	1010	5.00
1943	DMSP_f11	93210	094500	EAS	1465	2083	2.70
631	GMS_4	93210	102700	EAS	1376	1010	5.00
1794	NOAA_12	93210	104427	EAS	409	1730	4.00
636	GMS_4	93210	113400	EAS	1376	1010	5.00
641	GMS_4	93210	123500	EAS	1376	1010	5.00
646	GMS_4	93210	133500	EAS	1376	1010	5.00
651	GMS_4	93210	143400	EAS	1376	1010	5.00
1938	DMSP_f10	93210	144600	EAS	1465	865	2.70
656	GMS_4	93210	153400	EAS	1376	1010	5.00
661	GMS_4	93210	162700	EAS	1376	1010	5.00
666	GMS_4	93210	173400	EAS	1376	1010	5.00
1561	NOAA_11	93210	180817	EAS	409	1147	4.00
671	GMS_4	93210	183400	EAS	1376	1010	5.00
676	GMS_4	93210	193400	EAS	1376	1010	5.00
1569	NOAA_11	93210	194628	EAS	409	2034	4.00
681	GMS_4	93210	203400	EAS	1376	1010	5.00
2036	DMSP_f11	93210	211300	EAS	1465	2395	2.70
1810	NOAA_12	93210	213009	EAS	409	1005	4.00
686	GMS_4	93210	213500	EAS	1376	1010	5.00
691	GMS_4	93210	222700	EAS	1376	1010	5.00
1818	NOAA_12	93210	230728	EAS	409	1988	4.00
696	GMS_4	93210	233300	EAS	1376	1010	5.00
701	GMS_4	93211	003500	EAS	1376	1010	5.00
706	GMS_4	93211	013500	EAS	1376	1010	5.00
711	GMS_4	93211	023500	EAS	1376	1010	5.00
716	GMS_4	93211	033400	EAS	1376	1010	5.00
721	GMS_4	93211	042800	EAS	1376	1010	5.00

ENTRY	SATELLITE	DATE	TIME	ROI	ELES	LINES	RESLN
1577	NOAA_11	93211	052928	EAS	409	1394	4.00
726	GMS_4	93211	053400	EAS	1376	1010	5.00
731	GMS_4	93211	063500	EAS	1376	1010	5.00
1585	NOAA_11	93211	070617	EAS	409	1787	4.00
736	GMS_4	93211	073500	EAS	1376	1010	5.00
2041	DMSP_f11	93211	075600	EAS	1465	1553	2.70
741	GMS_4	93211	083500	EAS	1376	1010	5.00
1834	NOAA_12	93211	084704	EAS	409	1205	4.00
81	DMSP_f11	93211	093200	EAS	1465	2214	2.70
746	GMS_4	93211	093500	EAS	1376	1010	5.00
1948	DMSP_f10	93211	100000	EAS	1465	654	2.70
751	GMS_4	93211	102700	EAS	1376	1010	5.00
756	GMS_4	93211	113400	EAS	1376	1010	5.00
1826	NOAA_12	93211	120359	EAS	409	538	4.00
881	GMS_4	93211	133500	EAS	1376	1010	5.00
886	GMS_4	93211	143400	EAS	1376	1010	5.00
891	GMS_4	93211	153400	EAS	1376	1010	5.00
896	GMS_4	93211	162700	EAS	1376	1010	5.00
901	GMS_4	93211	173400	EAS	1376	1010	5.00
1593	NOAA_11	93211	175639	EAS	409	949	4.00
906	GMS_4	93211	183400	EAS	1376	1010	5.00
911	GMS_4	93211	193400	EAS	1376	1010	5.00
1601	NOAA_11	93211	193425	EAS	409	1998	4.00
916	GMS_4	93211	203400	EAS	1376	1010	5.00
1850	NOAA_12	93211	210947	EAS	409	680	4.00
921	GMS_4	93211	213500	EAS	1376	1010	5.00
926	GMS_4	93211	222700	EAS	1376	1010	5.00
1858	NOAA_12	93211	224618	EAS	409	1920	4.00
931	GMS_4	93211	233300	EAS	1376	1010	5.00
936	GMS_4	93212	003500	EAS	1376	1010	5.00
1953	DMSP_f10	93212	005400	EAS	1465	1886	2.70
941	GMS_4	93212	013500	EAS	1376	1010	5.00
946	GMS_4	93212	023500	EAS	1376	1010	5.00
951	GMS_4	93212	033500	EAS	1376	1010	5.00
956	GMS_4	93212	042800	EAS	1376	1010	5.00
1625	NOAA_11	93212	051736	EAS	409	1256	4.00
961	GMS_4	93212	053400	EAS	1376	1010	5.00
966	GMS_4	93212	063500	EAS	1376	1010	5.00
1609	NOAA_11	93212	065401	EAS	409	1876	4.00
971	GMS_4	93212	073600	EAS	1376	1010	5.00
1978	NOAA_12	93212	082602	EAS	409	901	4.00
976	GMS_4	93212	083500	EAS	1376	1010	5.00
981	GMS_4	93212	093500	EAS	1376	1010	5.00
986	GMS_4	93212	102700	EAS	1376	1010	5.00
991	GMS_4	93212	113400	EAS	1376	1010	5.00
1866	NOAA_12	93212	114213	EAS	409	975	4.00
996	GMS_4	93212	123500	EAS	1376	1010	5.00
1001	GMS_4	93212	133500	EAS	1376	1010	5.00
1006	GMS_4	93212	143400	EAS	1376	1010	5.00
1011	GMS_4	93212	153400	EAS	1376	1010	5.00
1016	GMS_4	93212	162700	EAS	1376	1010	5.00
1021	GMS_4	93212	173400	EAS	1376	1010	5.00

ENTRY	SATELLITE	DATE	TIME	ROI	ELES	LINES	RESLN
1649	NOAA_11	93212	174506	EAS	409	760	4.00
1026	GMS_4	93212	183400	EAS	1376	1010	5.00
1657	NOAA_11	93212	192225	EAS	409	1956	4.00
1031	GMS_4	93212	193400	EAS	1376	1010	5.00
1036	GMS_4	93212	203400	EAS	1376	1010	5.00
1041	GMS_4	93212	213500	EAS	1376	1010	5.00
1882	NOAA_12	93212	222514	EAS	409	1834	4.00
1046	GMS_4	93212	222700	EAS	1376	1010	5.00
1051	GMS_4	93212	233300	EAS	1376	1010	5.00

Appendix B

Archive Data Format Descriptions

By Data Processing Level

Level 1: Satellite Image Files

Satellite image filenames as they appear on tape have the following naming convention:

SSS_CCC_ROI_DDD_HH.Tif

where

SSS - Satellite identifier:

F10	DMSP F-10
F11	DMSP F-11
N11	NOAA-11
N12	NOAA-12
G04	GMS (Japan)

CCC - spectral channel identifier

ROI - Region of Interest:

EAS for East Asia Area

DDD - Julian day for which the image data are valid

HH - UTC hour of image data

Tif - TIF file format

File and Record Structure

All image files contain fixed-length records. The number of lines and number of elements in an image file are contained in the Related Entries (RE) SDB information file that is provided with the tape, under the heading of SATIMG:

NUM_LINES	Number of image data lines in the file.
ELEM_PER_LINE	Number of elements (pixels) per line.
BYTES_PER_ELEMENT	Number of bytes per pixel. This number is 1 for all SERCAA imager sensor data.

Image file data are stored in Tagged Image File Format (TIF), therefore an alternative way to determine image dimensions is to read the TIF header and examine the width and height fields.

Image pixel values represent either counts or albedo for visible data, and brightness temperatures for thermal infrared data. Table B-1 summarizes the attributes of the SERCAA image data values.

Table B-1 Satellite image characteristics

Satellite ID (SSS)	Spectral Channel (CCC)	Channel Type	Wavelength Band	Physical Value
F10 or F11	001	Visible	0.4 - 1.1 µm	Counts[1]
	002	Long-Wave IR	10 -12 µm	Brightness Temp.[2]
N11 or N12	001	Visible	0.63 µm	Albedo[3]
	002	Near-IR	0.86 µm	Albedo
	003	Mid-Wave IR	3.7 µm	Brightness Temp.
	004	Long-Wave IR	10.7 µm	Brightness Temp.
	005	Long-Wave IR	11.8 µm	Brightness Temp.
G04	001	Visible	0.55 - 0.75 µm	Counts
	002	Long-Wave IR	10.2 - 11.2 µm	Brightness Temp.

[1] Visible counts range from 0 - 255. High counts denote highly reflective surfaces and low counts denote poorly reflective surfaces.

[2] Brightness temperatures are byte-encoded such that the range 0 - 255 corresponds to the temperature range 327.5 K to 200.0 K. The relation between byte values and temperature is linear over this range; the conversion from byte value B to brightness temperature T is given by the relation:

$$T = -0.5\,B + 327.5.$$

[3] Albedo values are byte-encoded such that the range 0 - 255 corresponds to the albedo range 0 - 100%. The relation between byte values and percent albedo is linear; the conversion from byte value B to percent albedo A is given by the relation

$$A = 0.392\,B.$$

Level 1: **Latitude-Longitude File**

Latitude-longitude filenames as they appear on tape have the following naming convention:

SSS_LAT_ROI_DDD_HH.Dat

where

SSS - Satellite identifier:

F10	DMSP F-10
F11	DMSP F-11
N11	NOAA-11
N12	NOAA-12
G04	GMS (Japan)

LAT - a constant that identifies the file as a latitude-longitude file
ROI - Region of Interest for which the latitude-longitude file is valid:
EAS for East Asia Area
DDD - Julian day of satellite data for which the Earth locations are valid
HH - UTC hour of the satellite data for which the Earth locations are valid

File and Record Structure

Latitude-longitude Earth location files contain fixed-length records, the number and size of which depend on both the size of the corresponding image files and the satellite type. There is always one latitude-longitude record corresponding to each satellite image file record, where a satellite image file record contains one image scan line of information. However, latitude-longitude data are subsampled, relative to the sensor data, along a scan line. There is one latitude-longitude pair for every n^{th} image pixel, where n is a function of satellite. A linear interpolation is used to compute Earth location for intermediate pixels between latitude-longitude reference points.

The information necessary for interpreting a latitude-longitude file record is contained in the Related Entries (RE) SDB information file provided with the tape, under the heading of LATLON:

LL_REC_LEN	Record length in bytes.
LL_LINE_INTERVAL	The number of image file records per lat-lon record. For the July 1993 data set this number is always 1.
LL_ELEM_INTERVAL	The subsampling rate of lat-lon information relative to the corresponding satellite data. For example, if LL_ELEM_INTERVAL = 40, there is one latitude-longitude pair for every 40th image pixel in the scan line (i.e., for pixels 1, 41, 81, ...). Linear interpolation is required to retrieve Earth location information for intermediate pixels 2-40, 42-80, ...
LL_ELEM_PER_LINE	This is the number of latitude-longitude elements per latitude-longitude file record.

A latitude-longitude file data element is a 4-byte structure that contains the scaled latitude and longitude for a given pixel. Thus the length of a latitude-longitude file record in bytes is given by:

LL_REC_LEN = 4 * LL_ELEM_PER_LINE

where the 4 bytes consist of two 16-bit integer variables: LONG and LAT. The storage convention is as follows:

LONG	Pixel longitude * 128. To obtain the floating-point longitude, FLONG = LONG / 128. Longitude range is -180° to 180°, positive east.
LAT	Pixel latitude * 128. to obtain floating-point latitude, FLAT = LAT / 128. Latitude range is -90° to 90°, positive north.

Level 1: **Angles File**

The angles filenames as they appear on tape have the following naming convention:

SSS_ANG_ROI_DDD_HH.Dat

where

SSS - Satellite identifier:

F10	DMSP F-10
F11	DMSP F-11
N11	NOAA-11
N12	NOAA-12
G04	GMS-4 (Japan)

ANG - a constant that identifies the file as an angles file
ROI - Region of Interest for which the angles file is valid:
EAS for East Asia Area
DDD - Julian day of satellite data for which the angles are valid
HH - UTC hour of the satellite data for which the angles are valid

File and Record Structure

Angle files contain fixed-length records, the number and size of which depend on both the size of the corresponding image files and the satellite type. There is always one angles record corresponding to each satellite image file record, where a satellite image file record contains one image scan line of information. However, angle data are subsampled, relative to the sensor data, along a scan line. There is one set of angles for every n^{th} image pixel, where n is a function of satellite. A linear interpolation is used to compute angle values for intermediate pixels between angle reference points.

The information necessary for interpreting an angles file record is contained in the Related Entries (RE) SDB information file provided with the tape, under the heading of ANGLES:

ANG_REC_LEN	Record length in bytes.
ANG_LINE_INTERVAL	The number of image file records per angles record. This number is almost always 1.
ANG_ELEM_INTERVAL	The subsampling rate of angles information relative to the corresponding satellite image. For example, if ANG_ELEM_INTERVAL = 8, there is one set of angles valid for every eighth image pixel in the scan line (i.e., for pixels 1, 9, 17, 25, ...). Linear interpolation is required to retrieve angles information for intermediate pixels 2-8, 10-16, 18-24, ...
ANG_ELEM_PER_LINE	This is the number of angles elements per angles file record.

An angles file data element is a 12-byte structure containing three angles that define the satellite and solar viewing geometry for a given pixel. Thus the length of an angles file record in bytes is given by:

ANG_REC_LEN = 12 * ANG_ELEM_PER_LINE

where the 12 bytes consist of three 32-bit floating-point variables: SATZEN, SOLZEN, and AZIMUTH corresponding to the satellite zenith, the solar zenith, and the satellite/solar azimuth angles respectively (Figure B-1). Note: Angle files were generated on a VMS computer. To interpret these floating-point numbers on a UNIX machine it is necessary to convert from VMS to IEEE floating-point formats. Most UNIX operating systems provide a utility to perform this conversion. Angle measurement conventions are as follows:

SATZEN	Scene satellite zenith angle, 0^{o} - 90^{o}.
SOLZEN	Scene solar zenith angle, 0^{o} - 180^{o}.
AZIMUTH	Relative angle between the solar and satellite azimuth angles, 0^{o} - 359^{o}. When AZIMUTH = 0^{o}, the sun is directly behind the satellite (i.e., the viewed point, the satellite, and the sun are collinear). When AZIMUTH = 180^{o}, the satellite is looking directly into the sun (the satellite squints to compensate).

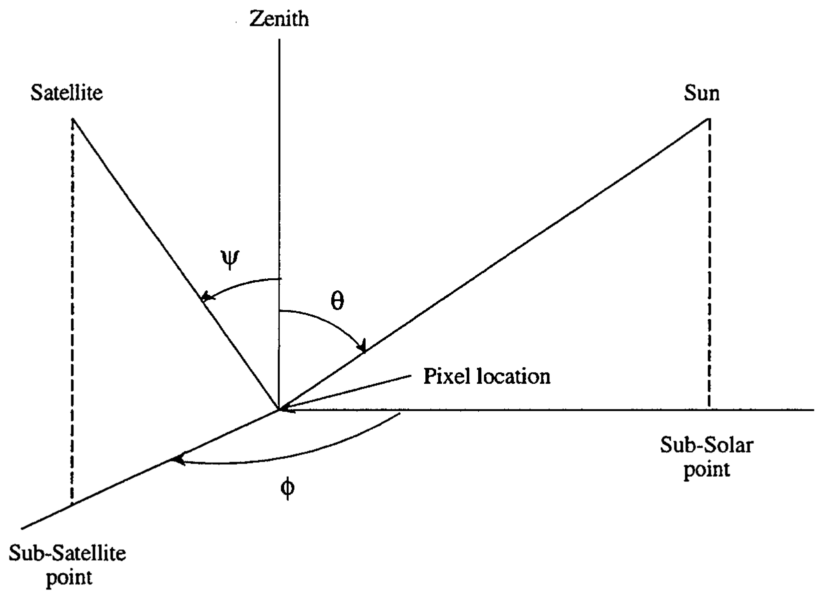


Figure B-1. Satellite-Earth-Solar Geometry (after Taylor and Stowe, 1984)

ψ - satellite zenith angle

θ - solar zenith angle

ϕ - sun-satellite azimuth angle

Level 2: **Nephanalysis Products**

Nephanalysis products are stored as bit-encoded byte values known as MCF (cloud Mask and Confidence Flag). MCF filenames as they appear on tape have the following naming convention:

SSS_MCF_ROI_DDD_HH.Dat or .Tif

where

SSS - Satellite identifier:

F10 DMSP F-10
F11 DMSP F-11
N11 NOAA-11
N12 NOAA-12
G04 GMS-4 (Japan)

MCF - a constant that identifies the file as an MCF file
ROI - Region of Interest for which the product is valid:
EAS for East Asia Area
DDD - Julian day for which the product is valid
HH - UTC hour for which the product is valid
Dat - Raw product file format
Tif - TIF file format

File and Record Structure

Level 2 processing is performed on square arrays of image pixels, therefore the size of the resultant MCF product files is an integral number of the analysis array size. MCF files contain fixed-length records, the number and size of which depends on both the size of the corresponding image files and the satellite type. The following table specifies how to determine the record size and number of records in an MCF file. Let NCOLS and NROWS be the number of columns and rows, respectively, in the corresponding satellite image file; then:

If the image satellite id is:	Then the MCF file record size is:	And the number of lines is:
DMSP F10 or F11	NCOLS - MOD(NCOLS, 16)	NROWS - MOD(NROWS, 16)
NOAA 11 or 12	NCOLS - MOD(NCOLS, 32)	NROWS - MOD(NROWS, 32)
GMS 4	See Associated RE File or TIF Header	See Associated RE File or TIF Header

where MOD is the FORTRAN modulus function (e.g., if an F10 pass has 1465 columns per scan line, then the MCF record size is 1456). The MCF file is stored in Tagged Image File Format (TIF), therefore an alternative way to determine file dimensions is to read the TIF header and examine the width and height fields.

The format of an MCF file is the same regardless of the satellite platform it was derived from. The first byte of the first record of the MCF file corresponds to the first byte of the first record in the corresponding image data file. Across each scan line there is a one-to-one correspondence between the image and MCF files out to the number of bytes computed above for each record. As can be seen in the above table, the MCF and image file sizes are not always the same. However, the two files are always aligned with respect to the upper-left corner of each.

There is one 8-bit MCF byte per analyzed image pixel. MCF bytes are bit-packed according to the following convention:

Bit 0 (least significant) is the cloud/no-cloud bit. If bit 0 is off, the corresponding image pixel is clear; if bit 0 is on, it is completely cloudy.

Bit 1 is the low cloud bit. If bit 1 is on, the pixel contains low cloud as determined by an appropriate spectral (or other) signature test.

Bit 2 is the thin cirrus cloud bit. If bit 2 is on, the pixel contains cirrus as determined by an appropriate spectral (or other) signature test.

Bit 3 is the cumulonimbus bit. If bit 3 is on, the pixel contains thunderstorm clouds.

Bit 4 is the partly cloudy bit. If bit 4 is on, the pixel is partly cloudy. If bit 4 is on, bit 0 is off. DMSP data are used exclusively to determine partly cloud conditions.

Bit 5 is the bad data bit. It is set whenever satellite data are missing or unreliable. If set, all other bits should be ignored.

Bits 6 and 7 contain the confidence level attached to the accuracy of the cloud/no-cloud decision for the corresponding cloudy image pixel. Confidence levels are rated as 0 for missing data, 1 for low confidence, 2 for mid-level confidence, and 3 for high confidence.

Low cloud, thin cirrus, and cumulonimbus conditions are always associated with completely cloudy conditions (i.e., bit 0 will always be on in the presence of one or more of these conditions). Cloud level and cloud type are not detected under partly cloudy conditions (i.e., if bit 4 is on, bits 1 through 3 will be off).

Example:

MCF byte	1	1	0	0	0	1	0	1	(C5 in hex)
bit position	7	6	5	4	3	2	1	0	

The corresponding image pixel is classified as cloud covered (bit 0) with thin cirrus (bit 2) that has been detected with a high level of confidence (bits 6 and 7).

Level 3: Layered Product

The layered product filename as it appears on tape has the following naming convention:

SAT_LYR_ROI_DDD_HH.DAT

where:

SAT - Satellite identifier:
- F10 DMSP F-10
- F11 DMSP F-11
- N11 NOAA-11
- N12 NOAA-12
- G04 GMS-4 (Japan)

LYR is a constant that denotes the file is a layered product

ROI - Region of Interest:
- EAS for the DNA East Asia Area (EASA)

DDD - Julian day

HH - GMT hour

File Structure

The layered product file contains 86175 (225 rows x 383 columns) record structures, each 55 bytes in length.

Record Structure

Each record contains data values valid for one grid point within a 383 (rows) by 225 (columns) two-dimensional grid. The grid is superimposed on a hemispheric secant polar stereographic map projection. Grid resolution is based on a whole mesh grid spacing of 381 km at 60 degrees latitude and nested grids are defined in terms of the number of grid cells that fit within a whole mesh grid. The layered product grid is a 1/16th mesh grid (i.e., 16 by 16 grid cells per whole mesh box.)

Table B-2 summarizes the contents of each record. Figure B-2 contains the C data structure that was used to create the data file.

Table B-2: Layered Product Record Structure

Field	Description	Units	Range	Missing or bad value	Byte length
1	Absolute 16th-mesh row number (i)		1-1024		2
2	Absolute 16th-mesh column number (j)		1-1024		2
3	SDB IR entry number			0	2
4	Julian day (ddd)				4
5	UTC (hhmm)		0-2359		2
6-9	Cloud temperature variance for each layer	GS*100			8
10	# pixels in grid box			0	2
11-14	# pixels in each layer			0	8
15-18	Cloud top temperature for each layer	GS*100			4
19-22	Cloud type for each layer		0-1		4
23-26	# low cloud pixels in each layer				4
27-30	# thin cirrus pixels in each layer				4
31-34	# precipitating-cloud pixels in each layer				4
35	Sunrise time		0-235		1
36	Sunset time		0-235		1
37	Satellite platform ID				1
38	# data dropouts in grid box				1
39	# partially cloud-filled pixels				1

```
/* Layering output structure

   Daniel Peduzzi (AER) 9/27/94
     structure content by Robert P. d'Entremont (AER) 9/1994
*/

#ifndef NCLASSES
#  define NCLASSES (4)
#endif

#ifndef _LAYER_OUTPUT
#define _LAYER_OUTPUT

#define BYTE unsigned char

typedef  struct {

  short i;                          /* 16th-mesh absolute row (1-1024)              */
  short j;                          /* 16th-mesh absolute column (1-1024)           */

  short sdb_ir_entry;               /* SDB entry number corresponding to IR data    */
  int yyddd;                        /* Sensor data Julian day                       */
  short hhmm;                       /* Sensor data valid time (UTC) hhmm            */

  short layer_var[NCLASSES];        /* Temperature variance*100 for cloud layer i   */
  short num_pixels;                 /* Total # of pixels in 16th-mesh box           */
  short n_layer_pix[NCLASSES];      /* Total # pixels in layer i                    */
  BYTE meantemp[NCLASSES];          /* Mean cloud top temperature for layer i       */
  BYTE cloud_type[NCLASSES];        /* Cloud type for layer i (1 or 2)              */
  BYTE low_cloud[NCLASSES];         /* # low cloud pixels in layer i                */
  BYTE thin_cirrus[NCLASSES];       /* # thin cirrus pixels in layer i              */
  BYTE precip[NCLASSES];            /* # precipitating-cloud pixels in layer i      */

  BYTE sunrise;                     /* Sunrise time (UTC) (0-235)                   */
  BYTE sunset;                      /* Sunset time (UTC) (0-235)                    */
  BYTE vid;                         /* Satellite vehicle (platform) ID              */
  BYTE dropouts;                    /* Total # of data dropouts in 16th-mesh box    */
  BYTE partial;                     /* Total # of partially-cloud-filled pixels     */

} LAYER_OUTPUT;

#undef BYTE

#endif
```

Figure B-2: Level 3 data structure

Level 4: Integrated Product

The integrated product filename as it appears on tape has the following naming convention:

ALL_IAN_ROI_DDD_HH.Dat

where

ALL and IAN are constants (Integrated ANalysis from ALL sensors)
ROI - Region of Interest for which the product is valid
Possible values:
EAS for the East Asia Area
DDD - Julian day for which the integrated product is valid
HH - GMT hour for which the integrated product is valid

File Structure

The integrated product file contains 86,175 records (225 columns by 383 rows), each 64 bytes in length.

Record Structure

Each record contains data values valid for one grid point within a 383 (rows) X 225 (columns) 2-D grid. The grid is superimposed on a hemispheric secant polar stereographic map projection. Grid resolution is based on a whole-mesh grid spacing of 381 km at 60^{o} latitude and nested grids are defined in terms of the number of grid cells that fit within a whole mesh grid. The integrated product grid is a $1/16^{th}$ mesh grid (i.e., 16 X 16 cells per whole mesh box).

Table B-3 summarizes the contents of each record. All values are 16-bit integers. Figure B-3 contains the C data structure used to create the output file.

Table B-3. Integrated Product Record Structure*

Field	Description	Units	Range	Missing or bad value	Comments
1	Absolute 16th-mesh column number (i)		227 - 451		
2	Absolute 16th-mesh row number (j)		13 - 395		
3	Number of cloud layers in (i,j)		0 - 4	-999	
4	Total cloud fraction for (i,j)	Percent	0 - 100	-999	
5-8	Cloud fraction by layer for (i,j)	Percent	0 - 100	-999	
9-12	Cloud top temperature by layer	K*10	2000-3275	-999	
13-16	Cloud top height by layer	Meters	0-13500	-999	
17-20	Cloud type by layer		0 - 9	-999	See Table B-4
21	Total cloud fraction error for (i,j)	Percent	0 - 100	-999	
22-25	Layer cloud fraction error for (i,j)	Percent	0 - 100	-999	
26-29	Layer confidence flags for (i,j)	Flag*10	10 - 30	-999	Discrete values for low to high confidence
30-32	Database entry numbers for input satellite analyses				Corresponds to directory names on tar tape

*all values are 16-bit integers

Table B-4. Cloud Type Codes

Cloud Type Code	Cloud Type
0	No Cloud
1	Cirrus
2	Cirrostratus
3	Altocumulus
4	Altostratus
5	Stratocumulus
6	Stratus
7	Cumulus
8	Cumulonimbus
9	Nimbostratus

```
/* EASA definitions */

#define NLINE 383
#define NCOL 225
#define NLYRS 4


#define MIN_I 227
#define MIN_J 13

typedef unsigned char byte;

/* integration output structure */

typedef struct {
   short i;                    /* absolute 16th mesh coord        */
   short j;
   short nlayers;              /* number of layers                */
   short fraction;             /* total cloud fraction            */

   short lyr_frc[NLYRS];       /* layer cloud fraction            */
   short t_cld[NLYRS];         /* layer cloud top temp (K*10)     */
   short z_cld[NLYRS];         /* layer cloud top height (m)      */
   short cld_typ[NLYRS];       /* layer cloud type                */

   short error;                /* total cloud amount error        */
   short lyr_err[NLYRS];       /* layer cloud amount error        */
   short conf[NLYRS];          /* layer confidence measure        */
   short sdb_entry[3];         /* input entry number(s)           */
} INTEGRATION;
```

Figure B-3: Integration output data structure

Appendix C

Data Extraction Guide

****SERCAA DATA SET RELEASE TO DNA****

****
What should I have ?

DNA_RELEASE.TXT	This document.
(2) 8 mm D8-112 tapes	One tape, labeled DNA JUL93 IA, contains the SERCAA Integrated Analysis (SIA) data files. The other tape, labeled DNA JUL93 ENTRIES contains the Related Entry (RE) data (which consists of Satellite, Latitude/Longitude, Angles(Geometry) and Product(cloud mask) data files.

****
What type of tape drive was used ?

A SUN Exabyte EXB-8500 8 mm tape drive recording in high density mode (5 gig).

****
What utility was used to create the release tapes ?

The data were placed on the tapes using a SUN SPARC II running SUN OS 4.1.2. The following tar command syntax was used:

```
sun% tar cvBf /dev/nrst8 somedirectory
```

****
How are the data arranged on the release tape ?

The data on the SIA tape are contained in 10 tar files. Each of these tar files represents a directory that contains all the SIA data for a particular day (day 93203 through day 93212). Each directory name follows the convention:

CYYJJJ

where:

C = century (9 for 19XX)
YY = year
JJJ = Julian day

A SIA file and SIA SDB information file exists for each hour that an analysis was performed. Each SIA file has been named using the following convention:

Positions 1-4 Platform:

all_ = All satellite platforms are used to create a SIA.

Positions 5-8 Type of file:

ian_ = integrated analysis file
sdb_ = SERCAA data base (SDB)
information file

Positions 9-12 Region of interest:
(Given in 16th-mesh coordinates)
eas_ = East Asia Area (EASA). (i,j) = (227,13) to (451,395)
can_ = Canada Area (CANA). (409,597) to (557,711)
cns_ = Central, Northern South America Area (CNSA). (413,877) to (651,1011)
emd_ = Eastern Mediterranean, Desert Area (EMDA). (731,353) to (863,505)

Positions 13-16 Julian day:
203_ = Julian day 203 etc. ...

Positions 17-18 Hour:
00 = SIA for hour 00 etc. ...

Positions 19-22 Extension:
.dat = raw-format file extension

Example: all_ian_eas_203_10.dat

The RE tape contains 320 files. Each of these tar files represent a directory that contains all the related data used as input to create at least one of the SIA data files. Each directory name follows the convention:

ENTRY/

where:
ENTRY = the SDB entry number

Each RE file has been named following these guidelines:

Positions 1-4 Platform:
n11_ = NOAA N_11
n12_ = NOAA N_12
f10_ = DMSP F_10
f11_ = DMSP F_11
g04_ = GMS-4

Positions 5-8 Type of file:
001_ = satellite data channel 1
002_ = satellite data channel 2
...
...
005_ = satellite data channel 5
lat_ = latlon data
ang_ = angles data
mcf_ = cloud mask data
sdb_ = SDB information file

Positions 9-12 Area of data:
eas_ = East Asia Area (EASA)
can_ = Canada Area (CANA)
cns_ = Central and Northern South America Area (CNSA)

emd_ = Easter Mediterranean, Desert Area (EMDA)

Positions 13-16 Julian day:

203_ = Julian day 203 etc. ...

Positions 17-18 Hour:

00 = hour of the data

Positions 19-22 Extension:

.dat = raw data
.tif = tif formatted data

Examples:

f10_001_eas_150_14.tif
f10_002_eas_150_14.tif
f10_lat_eas_150_14.tif
f10_ang_eas_150_14.tif
f10_mcf_eas_150_14.tif
f10_sdb_eas_150_14.tif

Refer to separate listing sheet labeled JUL93.IA.TAR.LIST for a listing of the IA tape contents. Run the provided script, "list_tar", to generate a listing of the RE tape.

****

What are related data items ?
What is the SDB entry number ?
What are related entries ?

The SDB registration process is a process that automatically places descriptive data items about a satellite scan into the SDB. The SDB registration process allocates a group of unique entry numbers to be used as place holders for all of the related data items for a given satellite scan. The related data items consists of satellite, latitude/longitude, angles (Geometry) and product(cloud mask) data. As an example, if a DMSP F_11 scan was to be registered in the SDB, the registration process would request for a group of five contiguous entry numbers(i.e. 1001-1005). These five entry numbers would be used as place holder for the following related data items:

1001	f11 visible channel
1002	f11 infrared channel
1003	latitude/longitude data
1004	angles(geometry) data
1005	product data

The "SDB entry number" is the first entry number of the group of entry numbers provided by the registration process. The first entry number is used to "key" into the related data items for that group. In the example provided above the SDB entry number would be 1001.

The release process uses the SDB entry number in each group to logically divide the data into separate directories (i.e. the directory name is first SDB entry number for each group of entry numbers). Using the example provided above the directory named "1001/" contains all the related data items for that group (i.e. the directory contains the data for entry 1001 through entry 1005).

To build a SIA it is necessary to use as input, related data items from one or more satellite scans and/or satellite platforms. The SDB entry number is used to keep track of all inputs to the SIA. The list of related entries are given as SDB entry numbers.

****

How do I get a particular SIA data set ?

You must use the UNIX tar utility to extract the data from the tape. By using the tar utility you may extract individual files or the entire directory. Use the JUL93.IA.TAR.LIST to determine where to position the tape and then use the appropriate tar command to extract the files you want. For example, if you want to extract all of the SIA data files from the first and second tar files, the following commands could be used:

```
% pwd
/users/smith
% mkdir data
% cd data
% tar xvf /dev/rst8 993203
% tar xvf /dev/rst8 993204
```

Upon completion all of the SIA data for day 203 would reside in directory /users/smith/data/993203 and all the SIA data for day 204 would reside in directory /users/smith/data/993204.

****

What is the SDB information file ?

The SDB information file is a text file containing selected SDB record items that help describe the actual data. The SIA SDB information file shows what data went into creating the SIA by listing the related entries. The RE SDB information file lists information about the satellite images, the latlon file, the angles file and the product file(s).

The following is an example SIA SDB information file:

```
[IA]
ZULU_YYJJJ:=93203                       : Year, Julian day of SIA
ZULU_HH:=10                             : Hour of SIA
ROI:=EAS                                : Region of Interest
NUM_RELATED_LAYER:=3                    : Number for related entries
RELATED_LAYER_1:= 4148                  : 1st related SDB entry number
RELATED_LAYER_2:= 7199                  : 2nd related SDB entry number
RELATED_LAYER_3:= 8988                  : 3d related SDB entry number
TDISK:=SDB_Int:
TDIR:=[SERCAA.DATA.993203]
FILE_IA_1:=ALL_IAN_EAS_203_08.Dat       : SIA file name
SDB_SET:=JUL93                          : Set identifier July of 1993
```

The following is an example RE SDB information file:

```
[SATIMG]
SAT_CODE:=16                            : Satellite code
ZULU_YYJJJ:=93203                       : Year, Julian day of scan
ZULU_HHMMSS:=82252                      : Time of scan
NUM_LINES:=1375                         : Number of lines
ELEM_PER_LINE:=409                      : Elements per line
```

BYTES_PER_ELEM:=1	: Bytes per element
7199:=AVH$005:[SERCAA.DATA.993203]N11_001_EAS_203_08.TIF	: Channel 1 file
7200:=AVH$005:[SERCAA.DATA.993203]N11_002_EAS_203_08.TIF	: Channel 2 file
7201:=AVH$005:[SERCAA.DATA.993203]N11_003_EAS_203_08.TIF	: Channel 3 file
7202:=AVH$005:[SERCAA.DATA.993203]N11_004_EAS_203_08.TIF	: Channel 4 file
7203:=AVH$005:[SERCAA.DATA.993203]N11_005_EAS_203_08.TIF	: Channel 5 file
[LATLON]	
LL_REC_LEN:=204	: Record length in bytes
LL_LINE_INTERVAL:=1	: Sub-sample line interval
LL_ELEM_INTERVAL:=8	: Sub-sample element interval
LL_ELEM_PER_LINE:=51	: Latlon pairs per line
LL_FILE:=AVH$005:[SERCAA.DATA.993203]N11_LAT_EAS_203_08.DAT	: latitude/longitude file
[ANGLES]	
ANG_REC_LEN:=612	: Record length in bytes
ANG_LINE_INTERVAL:=1	: Sub-sample line interval
ANG_ELEM_INTERVAL:=8	: Sub-sample element interval
ANG_ELEM_PER_LINE:=51	: Angles triplets per line
ANG_FILE:=AVH$005:[SERCAA.DATA.993203]N11_ANG_EAS_203_08.DAT	: Angles file
[PRODUCT]	
7206001:=sdb$prd:[SERCAA.DATA.993203]N11_MCF_SET_203_08.TIF	: Cloud mask file

****
How do I know which RE data went into a particular SIA ?

There are two ways to determine which RE data sets went into a particular SIA. The first way is reference the SIA SDB information file. Each "RELATED_LAYER" listed is a reference, by SDB entry number, to the RE data. Use the referred SDB entry number to retrieve the related data from the RE data tape.

For example, refer to the above SIA SDB information file. The "RELATED_LAYERED_1:=4148" line implies that SDB entry number 4148 and the related data items for entry 4148 (along with SDB entry numbers 7199 and 8988) were used to create "ALL_IAN_EAS_203_08.Dat".

The second way is to read the header information from the SIA file (Please refer to the DATA_DESCRIPTION).

****
How do I get the RE data files ?

Once you have examined the SIA SDB information file and you have identified the related entry numbers, you must use the UNIX tar utility to extract the data from the tape. By using the tar utility you may extract individual files or the entire directory. Use a tape contents list generated using the "list_tar" script to determine where to position the tape and then use the appropriate tar command to extract the files you want. For example, if you want to extract all of the RE data files from the first tar file, the following commands could be used:

```
% pwd
/users/smith
% mkdir data
% cd data
% tar xvf /dev/rst8 7199
```

Upon completion of this command all of the RE data related to SDB entry number 7199 would reside in directory /users/smith/data/7199.

****
For the following question please refer to the example SDB information files as needed.

****
What is the format of the satellite data and how do I access it?

The dimensions of the satellite data are defined by the three parameters, NUM_OF_LINES, ELEM_PER_LINE and BYTES_PER_ELEM . To access the data use the following logic.

If the file extension is ".dat"
then use the appropriate C or FORTRAN read statements.

If the file extension is ".tif"
then use a TIF reader or TIF library (you may view the images by using the public domain application, XV).

For a detailed explanation, refer to Appendix B.

****
What is the format of the latlon data and how do I access it?

The latlon data are sub-sampled. The dimensions are defined LL_LINE_INTERVAL, LL_ELEM_INTERVAL and LL_ELEM_PER_LINE. LL_ELEM_PER_LINE defines the number of longitude/latitude pairs per line. Each pair is four bytes (two bytes lon, two bytes lat). To access the data use the appropriate C or FORTRAN read statements.

For a detailed explanation, refer to Appendix B.

****
What is the format of the angles data and how do I access it?

The angles data are sub-sampled. The dimensions are defined by ANG_LINE_INTERVAL, ANG_ELEM_INTERVAL and ANG_ELEM_PER_LINE. ANG_ELEM_PER_LINE
defines the number of triplets (satellite-zenith/solar-zenith/azimuth) per line. Each item in the triplet is a float data type. To access the data use the appropriate C or FORTRAN read statements.

For a detailed explanation, refer to Appendix B.

P601

Data Save Documentation Report No. 4

ADVANCED GEOPHYSICAL ENVIRONMENT SIMULATION TECHNIQUES

Task 1: Satellite Data Sets for Worldwide Cloud Prediction

This data documentation report covers data set generation
for the DNA region of interest:

Central and Northern South America Area (CNSA)

for the period:

22-31 March 1994

Contract Number F19628-94-C-0106

issued by:

Electronic Systems Division
Air Force Systems Command
Hanscom AFB, MA 01731

Submitted by:

Atmospheric and Environmental Research, Inc.
840 Memorial Drive
Cambridge, MA 02139

3 March 1994

David B. Hogan
Gary B. Gustafson
Principal Investigators

1.0 Introduction

This Data Documentation Report provides a description of the fourth data save made in accordance with the revised statement of work for Satellite Data Sets for Worldwide Cloud Prediction Models. It is intended to provide a description of the data set, its format, how it was gathered and processed, and a description of the algorithms used to generate it. The data set consists of raw satellite data and analyzed products produced by the SERCAA cloud analysis algorithms. The period covered is 22-31 March 1994 for the DNA region of interest: Central and Northern South America (CNSA). This region covers the following (i,j) 16th mesh grid coordinates: 413,877 - 651,1011. All available data from those dates are included. These data were processed specifically for DNA using software developed from the SERCAA cloud analysis algorithms described by Gustafson et. al (1994). Substantial modifications were required to the Cloud Layering and Analysis Integration modules to accommodate the high volume of data included in this data set. Two tapes are provided, one with Level 1, 2 and 3 products and the second with Level 4.

2.0 Processing Environment

Satellite data processing for this data set used the SERCAA cloud analysis algorithms described by Gustafson et al. (1994). Multisource data from the DMSP F10 and F11, NOAA-11 and NOAA-12, and METEOSAT-3 satellites were used. Data sources were as follows: DMSP - National Geophysical Data Center (NGDC), Boulder, CO; NOAA - National Climatic Data Center (NCDC), Ashville, NC; METEOSAT - Phillips Laboratory direct readout ground station. All data were obtained by the Phillips Laboratory and were received on tape in various formats. All data processing was performed on the Air Force Interactive Meteorological System (AIMS) at the Phillips Laboratory. The SERCAA cloud analysis algorithms use four levels of data processing as summarized in Figure 1.

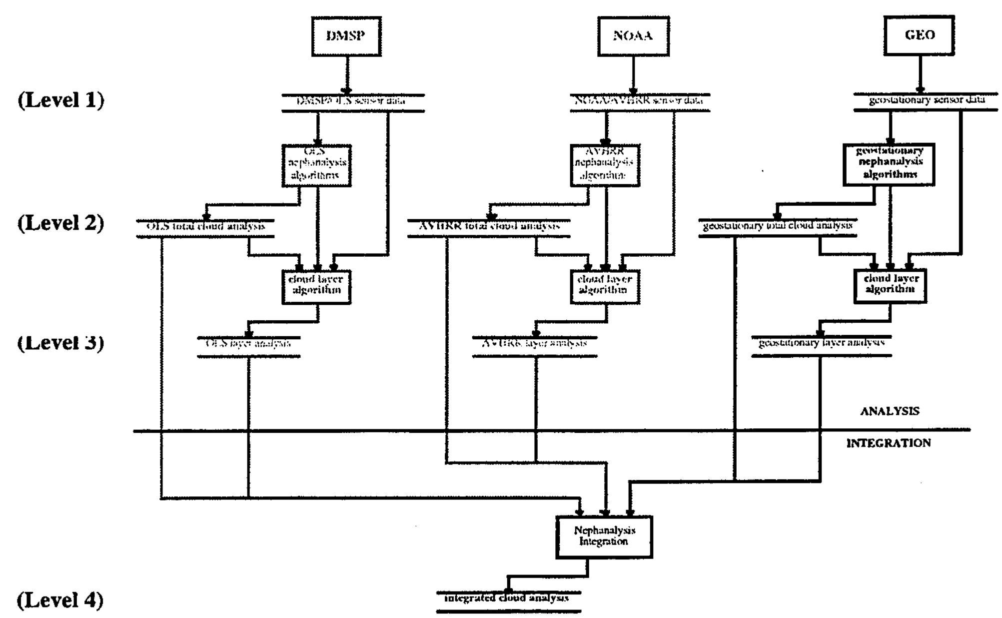


Figure 1 SERCAA data flow and processing levels

Level 1 processing consists of data ingest. Tape data are processed through separate ingest programs depending on the data source and format. All data are then stored in a standard format in the original satellite scan projection. The format consists of flat files where the number of elements correspond to the number of pixels in the satellite scan line and the number of rows corresponds to the number of scan lines. Data are maintained on AIMS through the SERCAA Database (SDB) management software. Level 1 data products consist of separate files for each sensor channel plus two additional files containing Earth location and satellite/solar geometry information. Satellite data characteristics are summarized in Table 1. In cases where visible and infrared channel resolution differ, the higher resolution data are subsampled to match the coarser resolution data (e.g., METEOSAT visible data are subsampled by a factor of two to match the IR data resolution). Earth location data consist of latitude-longitude pairs that are maintained at a subsampled resolution relative to the satellite data. For each sensor scan line, one latitude-longitude pair is provided for every n^{th} pixel, where n varies with satellite. Geometry information are also subsampled in the same ratio as the Earth location information and consist of three angles: satellite zenith, solar zenith, and sun-satellite azimuth. Ingest products are described more completely in Section 2 of Gustafson et al. (1994).

Table 1. Sensor Channel Data Attributes During SERCAA

Satellite	Sensor	Channel (μm)	Data Format	Resolution[1] (km)	Bits per Pixel[2]	Pixels per Scan Line
DMSP	OLS	0.40-1.10	counts	2.7	6	1464
		10.5-12.6	EBBT	2.7	8	1464
NOAA	AVHRR	0.58-0.68	percent albedo	4.0	10	409
		0.72-1.10	percent albedo	4.0	10	409
		3.55-3.93	EBBT	4.0	10	409
		10.3-11.3	EBBT	4.0	10	409
		11.5-12.5	EBBT	4.0	10	409
METEOSAT	VISSR	0.55-0.75	counts	2.5	8	5000
		10.5-12.6	EBBT	5.0	8	2500

[1]Sensor resolution at satellite subpoint that will provide global coverage.

[2]AVHRR radiance data are transmitted at 10-bit resolution, however, the SERCAA development system could only accommodate 8-bit brightness temperature data (although the full 10-bit resolution is used in the radiance-to-brightness-temperature transformation).

Level 2 processing consists of sensor-specific nephanalysis algorithms. Level 1 sensor data from DMSP, NOAA, and the METEOSAT geostationary satellites are processed through separate algorithms as indicated in Figure 1. Each time data from a new satellite pass are ingested, they are analyzed using the appropriate nephanalysis algorithm and results are placed in a Level 2 output file. One output file is generated for each nephanalysis run and nephanalysis results are stored in the original satellite scan projection with one byte of information for each pixel. Each byte is bit-packed according to the map in Table 2. For each set of Level 1 products generated from a satellite pass, one Level 2 product file is generated.

Table 2. Cloud Analysis Algorithm MCF File Bit Assignments

Bit	Assignment	Description
0	Cloud Mask	ON = Cloud-Filled OFF = Cloud-Free
1	Low Cloud	ON = Low Cloud Found
2	Thin Cirrus Cloud	ON = Thin Cirrus Cloud Found
3	Precipitating Cloud	ON = Precipitating Cloud Found
4	Partial Cloud	Only used by DMSP algorithm
5	Data Dropout	ON = Missing or Unreliable Data
6	Confidence	0 = Missing Data; 1 = Low;
7	Flag	2 = Middle; 3 = High

Level 3 processing uses Level 1 and 2 products as input to segment the cloudy regions into vertical cloud layers and to classify different cloud types. It also remaps the data from the individual satellite projections to the AFGWC standard polar stereographic map projection (Hoke et al., 1981) at 16th mesh grid resolution. The CNSA region of interest processed for the March 1994 data set have the following (i,j) 16th mesh grid coordinates: $413 \leq i \leq 651$, $877 \leq j \leq 1011$. Level 3 products are generated for each 16th mesh grid cell and contain the information in Table 3. A maximum of four cloud layers can be identified for each grid cell. One Level 3 file is created for each set of Level 1 and 2 products. All Level 1, 2, and 3 products associated with a single satellite pass are related through SDB and are provided on the DNA tapes as a set. Note that for the CNSA region, all Level 3 files are a fixed size of 239x135 grid cells.

Table 3. Cloud Typing and Layering Output

Parameter	Description
i	16th mesh i coordinate for Grid Cell
j	16th mesh j coordinate for Grid Cell
sdb_ir_entry	SDB entry number of input IR sensor data
ddd	Sensor data Julian date
hhmm	Sensor data valid time (UTC)
layer_var(4)	Cloud top IR variance of pixels in each layer
num_pixels	Total number of satellite pixels in 16th mesh grid cell
n_layer_pix(4)	Total number of pixels in each layer
meantemp(4)	Cloud top mean IR Temperature of pixels in each layer
cloud_type(4)	Cloud type of each layer
low_cloud(4)	Number of low cloud pixels in this layer detected by cloud analysis algorithm
thin_cirrus(4)	Number of thin cirrus pixels in this layer detected by cloud analysis algorithm
precip(4)	Number of precipitating cloud pixels in this layer detected by cloud analysis algorithm
sunrise	Local sunrise time (UTC)
sunset	Local sunset time (UTC)
vid	Satellite vehicle (platform) ID
dropouts	Number of bad data pixels in 16th mesh grid cell
partial	Number of partial cloud pixels detected by DMSP cloud analysis algorithm

Level 4 processing is a clock driven process with one new Level 4 integrated analysis performed each hour. Thus, integration is differentiated from the Level 1, 2, and 3 products that are event-driven (i.e., resulting from the ingest of a new satellite pass). The integration module operates on the most recent Level 3 gridded products available from each satellite source (i.e., NOAA, DMSP, METEOSAT). Like Level 3 products, the Level 4 output files conform to the AFGWC 16th mesh grid structure; output parameters for each grid cell are summarized in Table 4.

Table 4. Analysis Integration Processed Parameters

Parameter	Description
i	16th mesh i (column) coordinate
j	16th mesh j (row) coordinate
nlay	Number of Cloud Layers
cftot	Total Cloud Fraction
cf(4)	Layer Cloud Fraction
ctt(4)	Layer Cloud Top IR Temperature (K)
ctz(4)	Layer clout top height (m)
ity(4)	Layer Cloud Type
ecft	Estimated Error in Total Cloud Fraction
ecf(4)	Estimated Error in Layer Cloud Fraction
icf(4)	Analysis Confidence Flag Index For Each Layer
sdb(3)	SDB entry number of input analyses (NOAA, DMSP, METEOSAT)

3.0 Tape Format

All data for the March 1994 CNSA data save are contained on two 8 mm tapes written in UNIX tar format. The first tape, labeled: DNA MAR94 CNS ENTRIES (RE), contains all the Level 1-3 products. The second tape, labeled: DNA MAR94 CNS IA, contains all Level 4 products. The size of the combined Level 1, 2 and 3 products is approximately 1.8 Gbytes and the Level 4 products occupy 545 Mbytes. In addition to the two tapes, hard-copy listings of the contents of the Level 4 tape are also provided. The corresponding listing of the Level 1-3 tape is very large, so a UNIX script is provided to generate a listing at the user's site. It may be useful to place the listing file generated by the script into an edit program to scan and search it quickly. The listings are required to locate specific data sets on the tapes.

Level 1-3 products are generated for each new pass of satellite data received during the period of the data save. Appendix A contains a chronological list of each satellite pass used to produce the March 1994 data sets. All available data for the period covered were included; any gaps in the data list are due to either missing or bad data. DMSP data quality was improved over the previous, 1993 data sets. Although a few orbits of DMSP had to be dropped from the processing stream due to excessive bad or missing lines, there were no instances of the periodic drop-outs found in the earlier data sets. For data archiving purposes all Level 1-3 products associated with a given satellite pass were placed in a single directory and subsequently placed on tape as a single tar file. Thus the first tape contains a series of several hundred tar files; each file contains all Level 1-3 products associated with a single satellite pass. Level 4 files are grouped on the second

tape by day, thus for the March data save there are ten tar files on the Level 4 tape that each contain all Level 4 output files for each of the ten days 94081-94090 (22-31 March 1994). For each set of Level 1-3 products, and for each Level 4 file there is also an SDB Information File. These files contain descriptive metadata information extracted from the SERCAA Database that describe the relevant attributes of the SERCAA product files. For example, information files list the number of pixels in a scan line of satellite data and the number of scan lines in the file. Information on subsampling ratios for the Earth location and angles files are also contained there.

Detailed descriptions of the file formats used for each output level, and the associated information files, provided for the March 1994 save (Level 1, 2, 3, and 4) are provided in Appendix B. Appendix C provides a guide for extracting data sets from tape.

4.0 References

Gustafson, G.B., R.G. Isaacs, R.P. d'Entremont, J.M. Sparrow, T.M. Hamill, C. Grassotti, D.W. Johnson, C.P. Sarkisian, D.C. Peduzzi, B.T. Pearson, V.D. Jakabhazy, J.S. Belfiore, and A.S. Lisa, 1994: Support of Environmental Requirements for Cloud Analysis and Archive (SERCAA): algorithm descriptions. PL-TR-94-2114, Phillips Laboratory, Hanscom AFB, MA, ADA283240.

Hoke, J.E., J.L. Hayes, L.G. Renninger, 1981: Map projections and grid systems for meteorological applications. AFGWC-TN-79-003, Air Weather Service, Scott, AFB, IL.

Appendix A

Chronological List of Input Satellite Data

ENTRY	DATE	TIME	ROI	SATELLITE	ELES	LINES	RESLN
2018	94079	233557	CNS	NOAA_12	409	672	4.00
2034	94080	011640	CNS	NOAA_12	409	625	4.00
2058	94080	121046	CNS	NOAA_12	409	623	4.00
2082	94080	135107	CNS	NOAA_12	409	643	4.00
344	94080	150000	CNS	MET_3	896	512	5.00
349	94080	160000	CNS	MET_3	896	512	5.00
354	94080	170000	CNS	MET_3	896	512	5.00
359	94080	180000	CNS	MET_3	896	512	5.00
364	94080	190000	CNS	MET_3	896	512	5.00
369	94080	200000	CNS	MET_3	896	512	5.00
374	94080	210000	CNS	MET_3	896	512	5.00
379	94080	220000	CNS	MET_3	896	512	5.00
4432	94080	223700	CNS	DMSP	1465	736	2.70
384	94080	230000	CNS	MET_3	896	512	5.00
4442	94080	231800	CNS	DMSP	1465	700	2.70
1312	94081	000000	CNS	MET_3	896	512	5.00
2338	94081	005447	CNS	NOAA_12	409	655	4.00
388	94081	010000	CNS	MET_3	896	512	5.00
1152	94081	012600	CNS	DMSP	1465	701	2.70
392	94081	020000	CNS	MET_3	896	512	5.00
1696	94081	030600	CNS	DMSP	1465	688	2.70
396	94081	040000	CNS	MET_3	896	512	5.00
400	94081	070000	CNS	MET_3	896	512	5.00
404	94081	080000	CNS	MET_3	896	512	5.00
408	94081	090000	CNS	MET_3	896	512	5.00
412	94081	100000	CNS	MET_3	896	512	5.00
417	94081	110000	CNS	MET_3	896	512	5.00
2170	94081	114905	CNS	NOAA_12	409	595	4.00
422	94081	120000	CNS	MET_3	896	512	5.00
427	94081	130000	CNS	MET_3	896	512	5.00
2194	94081	132919	CNS	NOAA_12	409	665	4.00
432	94081	140000	CNS	MET_3	896	512	5.00
3200	94081	145400	CNS	DMSP	1465	713	2.70
437	94081	150000	CNS	MET_3	896	512	5.00
442	94081	160000	CNS	MET_3	896	512	5.00
3488	94081	163400	CNS	DMSP	1465	790	2.70
447	94081	170000	CNS	MET_3	896	512	5.00
452	94081	180000	CNS	MET_3	896	512	5.00
457	94081	190000	CNS	MET_3	896	512	5.00
462	94081	200000	CNS	MET_3	896	512	5.00
467	94081	210000	CNS	MET_3	896	512	5.00
4477	94081	212400	CNS	DMSP	1465	721	2.70
472	94081	220000	CNS	MET_3	896	512	5.00
2346	94081	225244	CNS	NOAA_12	409	572	4.00
477	94081	230000	CNS	MET_3	896	512	5.00
4487	94081	230500	CNS	DMSP	1465	710	2.70
1366	94082	000000	CNS	MET_3	896	512	5.00
2426	94082	003256	CNS	NOAA_12	409	689	4.00
481	94082	010000	CNS	MET_3	896	512	5.00

ENTRY	DATE	TIME	ROI	SATELLITE	ELES	LINES	RESLN
485	94082	020000	CNS	MET_3	896	512	5.00
489	94082	030000	CNS	MET_3	896	512	5.00
4412	94082	033400	CNS	DMSP	1465	754	2.70
493	94082	040000	CNS	MET_3	896	512	5.00
4497	94082	041400	CNS	DMSP	1465	605	2.70
497	94082	070000	CNS	MET_3	896	512	5.00
501	94082	080000	CNS	MET_3	896	512	5.00
12	94082	082501	CNS	NOAA_11	409	579	4.00
505	94082	100000	CNS	MET_3	896	512	5.00
4577	94082	100300	CNS	DMSP	1465	722	2.70
510	94082	110000	CNS	MET_3	896	512	5.00
2290	94082	112724	CNS	NOAA_12	409	583	4.00
515	94082	120000	CNS	MET_3	896	512	5.00
4587	94082	124400	CNS	DMSP	1465	808	2.70
520	94082	130000	CNS	MET_3	896	512	5.00
2314	94082	130738	CNS	NOAA_12	409	671	4.00
525	94082	140000	CNS	MET_3	896	512	5.00
530	94082	150000	CNS	MET_3	896	512	5.00
4512	94082	150300	CNS	DMSP	1465	725	2.70
535	94082	160000	CNS	MET_3	896	512	5.00
540	94082	170000	CNS	MET_3	896	512	5.00
545	94082	180000	CNS	MET_3	896	512	5.00
550	94082	190000	CNS	MET_3	896	512	5.00
555	94082	200000	CNS	MET_3	896	512	5.00
560	94082	210000	CNS	MET_3	896	512	5.00
311	94082	211351	CNS	NOAA_11	409	705	4.00
565	94082	220000	CNS	MET_3	896	512	5.00
2354	94082	223342	CNS	NOAA_12	409	180	4.00
68	94082	225533	CNS	NOAA_11	409	596	4.00
570	94082	230000	CNS	MET_3	896	512	5.00
4597	94082	235200	CNS	DMSP	1465	755	2.70
2367	94083	000000	CNS	MET_3	896	512	5.00
2434	94083	001119	CNS	NOAA_12	409	695	4.00
574	94083	010000	CNS	MET_3	896	512	5.00
2394	94083	015224	CNS	NOAA_12	409	590	4.00
578	94083	020000	CNS	MET_3	896	512	5.00
4537	94083	020200	CNS	DMSP	1465	785	2.70
582	94083	040000	CNS	MET_3	896	512	5.00
4547	94083	044300	CNS	DMSP	1465	635	2.70
586	94083	070000	CNS	MET_3	896	512	5.00
590	94083	080000	CNS	MET_3	896	512	5.00
105	94083	081247	CNS	NOAA_11	409	555	4.00

ENTRY	DATE	TIME	ROI	SATELLITE	ELES	LINES	RESLN
121	94083	095337	CNS	NOAA_11	409	672	4.00
594	94083	100000	CNS	MET_3	896	512	5.00
4617	94083	105000	CNS	DMSP	1465	714	2.70
599	94083	110000	CNS	MET_3	896	512	5.00
2418	94083	110552	CNS	NOAA_12	409	545	4.00
143	94083	113509	CNS	NOAA_11	409	440	4.00
604	94083	120000	CNS	MET_3	896	512	5.00
4627	94083	123100	CNS	DMSP	1465	781	2.70
3194	94083	124604	CNS	NOAA_12	409	664	4.00
609	94083	130000	CNS	MET_3	896	512	5.00
614	94083	140000	CNS	MET_3	896	512	5.00
3210	94083	142706	CNS	NOAA_12	409	344	4.00
619	94083	150000	CNS	MET_3	896	512	5.00
4562	94083	153100	CNS	DMSP	1465	762	2.70
624	94083	160000	CNS	MET_3	896	512	5.00
4572	94083	161100	CNS	DMSP	1465	729	2.70
629	94083	170000	CNS	MET_3	896	512	5.00
634	94083	180000	CNS	MET_3	896	512	5.00
639	94083	190000	CNS	MET_3	896	512	5.00
644	94083	200000	CNS	MET_3	896	512	5.00
649	94083	210000	CNS	MET_3	896	512	5.00
740	94083	210133	CNS	NOAA_11	409	699	4.00
654	94083	220000	CNS	MET_3	896	512	5.00
207	94083	224308	CNS	NOAA_11	409	605	4.00
659	94083	230000	CNS	MET_3	896	512	5.00
4737	94083	233900	CNS	DMSP	1465	786	2.70
3386	94083	234940	CNS	NOAA_12	409	687	4.00
2371	94084	000000	CNS	MET_3	896	512	5.00
4747	94084	002100	CNS	DMSP	1465	627	2.70
663	94084	010000	CNS	MET_3	896	512	5.00
3274	94084	013036	CNS	NOAA_12	409	604	4.00
667	94084	020000	CNS	MET_3	896	512	5.00
4662	94084	031100	CNS	DMSP	1465	683	2.70
671	94084	040000	CNS	MET_3	896	512	5.00
675	94084	070000	CNS	MET_3	896	512	5.00
679	94084	080000	CNS	MET_3	896	512	5.00
239	94084	080036	CNS	NOAA_11	409	526	4.00
263	94084	094121	CNS	NOAA_11	409	669	4.00
683	94084	100000	CNS	MET_3	896	512	5.00
688	94084	110000	CNS	MET_3	896	512	5.00
4757	94084	111800	CNS	DMSP	1465	749	2.70
279	94084	112236	CNS	NOAA_11	409	616	4.00
693	94084	120000	CNS	MET_3	896	512	5.00
3306	94084	122429	CNS	NOAA_12	409	643	4.00
698	94084	130000	CNS	MET_3	896	512	5.00
4682	94084	140000	CNS	DMSP	1465	719	2.70
703	94084	140000	CNS	MET_3	896	512	5.00

ENTRY	DATE	TIME	ROI	SATELLITE	ELES	LINES	RESLN
3322	94084	140502	CNS	NOAA_12	409	621	4.00
708	94084	150000	CNS	MET_3	896	512	5.00
713	94084	160000	CNS	MET_3	896	512	5.00
4692	94084	163900	CNS	DMSP	1465	799	2.70
718	94084	170000	CNS	MET_3	896	512	5.00
723	94084	180000	CNS	MET_3	896	512	5.00
728	94084	190000	CNS	MET_3	896	512	5.00
733	94084	200000	CNS	MET_3	896	512	5.00
756	94084	204915	CNS	NOAA_11	409	688	4.00
128	94084	210000	CNS	MET_3	896	512	5.00
256	94084	220000	CNS	MET_3	896	512	5.00
4782	94084	222500	CNS	DMSP	1465	807	2.70
764	94084	223040	CNS	NOAA_11	409	623	4.00
866	94084	230000	CNS	MET_3	896	512	5.00
3410	94084	232803	CNS	NOAA_12	409	660	4.00
2463	94085	000000	CNS	MET_3	896	512	5.00
4917	94085	000700	CNS	DMSP	1465	636	2.70
882	94085	010000	CNS	MET_3	896	512	5.00
3418	94085	010840	CNS	NOAA_12	409	636	4.00
870	94085	020000	CNS	MET_3	896	512	5.00
874	94085	040000	CNS	MET_3	896	512	5.00
4707	94085	042000	CNS	DMSP	1465	596	2.70
878	94085	070000	CNS	MET_3	896	512	5.00
886	94085	080000	CNS	MET_3	896	512	5.00
4842	94085	092400	CNS	DMSP	1465	666	2.70
788	94085	092905	CNS	NOAA_11	409	652	4.00
890	94085	100000	CNS	MET_3	896	512	5.00
895	94085	110000	CNS	MET_3	896	512	5.00
4847	94085	110500	CNS	DMSP	1465	707	2.70
812	94085	111006	CNS	NOAA_11	409	643	4.00
900	94085	120000	CNS	MET_3	896	512	5.00
2450	94085	120250	CNS	NOAA_12	409	611	4.00
905	94085	130000	CNS	MET_3	896	512	5.00
4717	94085	132800	CNS	DMSP	1465	682	2.70
2482	94085	134307	CNS	NOAA_12	409	654	4.00
910	94085	140000	CNS	MET_3	896	512	5.00
915	94085	150000	CNS	MET_3	896	512	5.00
4792	94085	150800	CNS	DMSP	1465	742	2.70
920	94085	160000	CNS	MET_3	896	512	5.00
925	94085	170000	CNS	MET_3	896	512	5.00
930	94085	180000	CNS	MET_3	896	512	5.00
943	94085	190000	CNS	MET_3	896	512	5.00
948	94085	200000	CNS	MET_3	896	512	5.00
1219	94085	203658	CNS	NOAA_11	409	669	4.00
953	94085	210000	CNS	MET_3	896	512	5.00
958	94085	220000	CNS	MET_3	896	512	5.00
4867	94085	221200	CNS	DMSP	1465	822	2.70

ENTRY	DATE	TIME	ROI	SATELLITE	ELES	LINES	RESLN
1235	94085	221811	CNS	NOAA_11	409	644	4.00
963	94085	230000	CNS	MET_3	896	512	5.00
2762	94085	230625	CNS	NOAA_12	409	615	4.00
2467	94086	000000	CNS	MET_3	896	512	5.00
2818	94086	004648	CNS	NOAA_12	409	666	4.00
4877	94086	005400	CNS	DMSP	1465	644	2.70
1007	94086	010000	CNS	MET_3	896	512	5.00
1019	94086	020000	CNS	MET_3	896	512	5.00
4817	94086	020700	CNS	DMSP	1465	781	2.70
1035	94086	040000	CNS	MET_3	896	512	5.00
4827	94086	044800	CNS	DMSP	1465	634	2.70
1047	94086	070000	CNS	MET_3	896	512	5.00
1059	94086	080000	CNS	MET_3	896	512	5.00
1063	94086	090000	CNS	MET_3	896	512	5.00
1002	94086	091647	CNS	NOAA_11	409	634	4.00
1075	94086	100000	CNS	MET_3	896	512	5.00
1042	94086	105740	CNS	NOAA_11	409	664	4.00
1080	94086	110000	CNS	MET_3	896	512	5.00
1093	94086	120000	CNS	MET_3	896	512	5.00
1106	94086	130000	CNS	MET_3	896	512	5.00
2594	94086	132121	CNS	NOAA_12	409	670	4.00
1111	94086	140000	CNS	MET_3	896	512	5.00
1124	94086	150000	CNS	MET_3	896	512	5.00
1137	94086	160000	CNS	MET_3	896	512	5.00
4937	94086	161600	CNS	DMSP	1465	733	2.70
1142	94086	170000	CNS	MET_3	896	512	5.00
1147	94086	180000	CNS	MET_3	896	512	5.00
1160	94086	190000	CNS	MET_3	896	512	5.00
1165	94086	200000	CNS	MET_3	896	512	5.00
1263	94086	202441	CNS	NOAA_11	409	643	4.00
1178	94086	210000	CNS	MET_3	896	512	5.00
1191	94086	220000	CNS	MET_3	896	512	5.00
1478	94086	220546	CNS	NOAA_11	409	661	4.00
2826	94086	224550	CNS	NOAA_12	409	423	4.00
1196	94086	230000	CNS	MET_3	896	512	5.00
2527	94087	000000	CNS	MET_3	896	512	5.00
2834	94087	002503	CNS	NOAA_12	409	690	4.00
5042	94087	004100	CNS	DMSP	1465	667	2.70
1256	94087	010000	CNS	MET_3	896	512	5.00
1268	94087	020000	CNS	MET_3	896	512	5.00
4962	94087	023600	CNS	DMSP	1465	708	2.70
1272	94087	030000	CNS	MET_3	896	512	5.00
1284	94087	040000	CNS	MET_3	896	512	5.00
1288	94087	070000	CNS	MET_3	896	512	5.00
1300	94087	083000	CNS	MET_3	896	512	5.00
1251	94087	090428	CNS	NOAA_11	409	614	4.00
1324	94087	100000	CNS	MET_3	896	512	5.00
1307	94087	104516	CNS	NOAA_11	409	676	4.00
1329	94087	110000	CNS	MET_3	896	512	5.00

ENTRY	DATE	TIME	ROI	SATELLITE	ELES	LINES	RESLN
5057	94087	113800	CNS	DMSP	1465	773	2.70
1334	94087	120000	CNS	MET_3	896	512	5.00
5067	94087	121900	CNS	DMSP	1465	745	2.70
2706	94087	125944	CNS	NOAA_12	409	670	4.00
1339	94087	130000	CNS	MET_3	896	512	5.00
1052	94087	140000	CNS	MET_3	896	512	5.00
4992	94087	140500	CNS	DMSP	1465	723	2.70
1183	94087	150000	CNS	MET_3	896	512	5.00
1277	94087	160000	CNS	MET_3	896	512	5.00
1352	94087	170000	CNS	MET_3	896	512	5.00
1357	94087	180000	CNS	MET_3	896	512	5.00
1362	94087	190000	CNS	MET_3	896	512	5.00
1510	94087	201224	CNS	NOAA_11	409	615	4.00
1372	94087	210000	CNS	MET_3	896	512	5.00
1526	94087	215320	CNS	NOAA_11	409	681	4.00
1377	94087	220000	CNS	MET_3	896	512	5.00
5087	94087	224600	CNS	DMSP	1465	740	2.70
1382	94087	230000	CNS	MET_3	896	512	5.00
5097	94087	232800	CNS	DMSP	1465	689	2.70
2531	94088	000000	CNS	MET_3	896	512	5.00
2842	94088	000323	CNS	NOAA_12	409	695	4.00
1386	94088	010000	CNS	MET_3	896	512	5.00
2778	94088	014427	CNS	NOAA_12	409	593	4.00
1390	94088	020000	CNS	MET_3	896	512	5.00
5112	94088	034400	CNS	DMSP	1465	735	2.70
1394	94088	040000	CNS	MET_3	896	512	5.00
1398	94088	070000	CNS	MET_3	896	512	5.00
1402	94088	080000	CNS	MET_3	896	512	5.00
1622	94088	085208	CNS	NOAA_11	409	602	4.00
1406	94088	100000	CNS	MET_3	896	512	5.00
5187	94088	102500	CNS	DMSP	1465	744	2.70
1646	94088	103255	CNS	NOAA_11	409	685	4.00
1411	94088	110000	CNS	MET_3	896	512	5.00
1416	94088	120000	CNS	MET_3	896	512	5.00
5197	94088	120600	CNS	DMSP	1465	735	2.70
2858	94088	123810	CNS	NOAA_12	409	657	4.00
1421	94088	130000	CNS	MET_3	896	512	5.00
1426	94088	140000	CNS	MET_3	896	512	5.00
2874	94088	141900	CNS	NOAA_12	409	505	4.00
5117	94088	143400	CNS	DMSP	1465	696	2.70
1431	94088	150000	CNS	MET_3	896	512	5.00
1436	94088	160000	CNS	MET_3	896	512	5.00
1441	94088	170000	CNS	MET_3	896	512	5.00
1446	94088	180000	CNS	MET_3	896	512	5.00
1451	94088	190000	CNS	MET_3	896	512	5.00
1706	94088	200015	CNS	NOAA_11	409	560	4.00
1461	94088	210000	CNS	MET_3	896	512	5.00
3450	94088	214057	CNS	NOAA_11	409	697	4.00

ENTRY	DATE	TIME	ROI	SATELLITE	ELES	LINES	RESLN
1466	94088	220000	CNS	MET_3	896	512	5.00
5212	94088	223300	CNS	DMSP	1465	729	2.70
1471	94088	230000	CNS	MET_3	896	512	5.00
5222	94088	231400	CNS	DMSP	1465	704	2.70
1738	94088	232252	CNS	NOAA_11	409	568	4.00
3090	94088	234145	CNS	NOAA_12	409	679	4.00
2559	94089	000000	CNS	MET_3	896	512	5.00
983	94089	010000	CNS	MET_3	896	512	5.00
3130	94089	012235	CNS	NOAA_12	409	613	4.00
987	94089	020000	CNS	MET_3	896	512	5.00
1491	94089	040000	CNS	MET_3	896	512	5.00
5162	94089	045300	CNS	DMSP	1465	628	2.70
1495	94089	070000	CNS	MET_3	896	512	5.00
1499	94089	080000	CNS	MET_3	896	512	5.00
1798	94089	083949	CNS	NOAA_11	409	596	4.00
1503	94089	090000	CNS	MET_3	896	512	5.00
1747	94089	100000	CNS	MET_3	896	512	5.00
5242	94089	101200	CNS	DMSP	1465	732	2.70
1940	94089	102037	CNS	NOAA_11	409	686	4.00
1760	94089	110000	CNS	MET_3	896	512	5.00
1765	94089	120000	CNS	MET_3	896	512	5.00
2970	94089	121634	CNS	NOAA_12	409	633	4.00
5252	94089	125300	CNS	DMSP	1465	746	2.70
1778	94089	130000	CNS	MET_3	896	512	5.00
2994	94089	135659	CNS	NOAA_12	409	635	4.00
1791	94089	140000	CNS	MET_3	896	512	5.00
1804	94089	150000	CNS	MET_3	896	512	5.00
5177	94089	154200	CNS	DMSP	1465	781	2.70
1809	94089	160000	CNS	MET_3	896	512	5.00
5272	94089	162100	CNS	DMSP	1465	733	2.70
1822	94089	170000	CNS	MET_3	896	512	5.00
1827	94089	180000	CNS	MET_3	896	512	5.00
1832	94089	190000	CNS	MET_3	896	512	5.00
3562	94089	194928	CNS	NOAA_11	409	334	4.00
1837	94089	200000	CNS	MET_3	896	512	5.00
1842	94089	210000	CNS	MET_3	896	512	5.00
3610	94089	212839	CNS	NOAA_11	409	703	4.00
1847	94089	220000	CNS	MET_3	896	512	5.00
1852	94089	230000	CNS	MET_3	896	512	5.00
5377	94089	230100	CNS	DMSP	1465	712	2.70
3506	94089	231026	CNS	NOAA_11	409	586	4.00
3154	94089	232009	CNS	NOAA_12	409	644	4.00

ENTRY	DATE	TIME	ROI	SATELLITE	ELES	LINES	RESLN
2563	94090	000000	CNS	MET_3	896	512	5.00
1856	94090	010000	CNS	MET_3	896	512	5.00
3170	94090	010040	CNS	NOAA_12	409	646	4.00
1860	94090	020000	CNS	MET_3	896	512	5.00
5297	94090	032100	CNS	DMSP	1465	667	2.70
1864	94090	040000	CNS	MET_3	896	512	5.00
1868	94090	043000	CNS	MET_3	896	512	5.00
1872	94090	070000	CNS	MET_3	896	512	5.00
1876	94090	080000	CNS	MET_3	896	512	5.00
3538	94090	082732	CNS	NOAA_11	409	581	4.00
1880	94090	100000	CNS	MET_3	896	512	5.00
3554	94090	100822	CNS	NOAA_11	409	682	4.00
5387	94090	105900	CNS	DMSP	1465	719	2.70
1885	94090	110000	CNS	MET_3	896	512	5.00
3570	94090	115018	CNS	NOAA_11	409	178	4.00
3082	94090	115453	CNS	NOAA_12	409	602	4.00
1890	94090	120000	CNS	MET_3	896	512	5.00
5397	94090	124000	CNS	DMSP	1465	798	2.70
1895	94090	130000	CNS	MET_3	896	512	5.00
3106	94090	133508	CNS	NOAA_12	409	661	4.00
1900	94090	140000	CNS	MET_3	896	512	5.00
5317	94090	141000	CNS	DMSP	1465	727	2.70
1905	94090	150000	CNS	MET_3	896	512	5.00
1910	94090	160000	CNS	MET_3	896	512	5.00
5327	94090	165000	CNS	DMSP	1465	723	2.70
1915	94090	170000	CNS	MET_3	896	512	5.00
1920	94090	180000	CNS	MET_3	896	512	5.00
1925	94090	190000	CNS	MET_3	896	512	5.00
1532	94090	200000	CNS	MET_3	896	512	5.00
1537	94090	210000	CNS	MET_3	896	512	5.00
3650	94090	211620	CNS	NOAA_11	409	706	4.00
1542	94090	220000	CNS	MET_3	896	512	5.00
3634	94090	225803	CNS	NOAA_11	409	595	4.00
3178	94090	225830	CNS	NOAA_12	409	591	4.00
1547	94090	230000	CNS	MET_3	896	512	5.00
5422	94090	234800	CNS	DMSP	1465	766	2.70
2911	94091	000000	CNS	MET_3	896	512	5.00
3426	94091	003849	CNS	NOAA_12	409	369	4.00
1551	94091	010000	CNS	MET_3	896	512	5.00
5432	94091	013000	CNS	DMSP	1465	615	2.70
1659	94091	020000	CNS	MET_3	896	512	5.00
1691	94091	043000	CNS	MET_3	896	512	5.00
1945	94091	070000	CNS	MET_3	896	512	5.00

ENTRY	DATE	TIME	ROI	SATELLITE	ELES	LINES	RESLN
1949	94091	080000	CNS	MET_3	896	512	5.00
1953	94091	100000	CNS	MET_3	896	512	5.00
1958	94091	110000	CNS	MET_3	896	512	5.00
1963	94091	120000	CNS	MET_3	896	512	5.00
1968	94091	130000	CNS	MET_3	896	512	5.00
1973	94091	140000	CNS	MET_3	896	512	5.00
1978	94091	150000	CNS	MET_3	896	512	5.00
1983	94091	160000	CNS	MET_3	896	512	5.00
1988	94091	170000	CNS	MET_3	896	512	5.00
1993	94091	180000	CNS	MET_3	896	512	5.00
1998	94091	190000	CNS	MET_3	896	512	5.00
2003	94091	200000	CNS	MET_3	896	512	5.00

Appendix B

Archive Data Format Descriptions

By Data Processing Level

Level 1: Satellite Image Files

Satellite image filenames as they appear on tape have the following naming convention:

SSS_CCC_ROI_DDD_HH.Tif

where

SSS - Satellite identifier:

F10 DMSP F-10
F11 DMSP F-11
N11 NOAA-11
N12 NOAA-12
M03 METEOSAT-3

CCC - spectral channel identifier

ROI - Region of Interest:

CNS for Central and Northern South America Area

DDD - Julian day for which the image data are valid

HH - UTC hour of image data

Tif - TIF file format

File and Record Structure

All image files contain fixed-length records. The number of lines and number of elements in an image file are contained in the Related Entries (RE) SDB information file that is provided with the tape, under the heading of SATIMG:

NUM_LINES	Number of image data lines in the file.
ELEM_PER_LINE	Number of elements (pixels) per line.
BYTES_PER_ELEMENT	Number of bytes per pixel. This number is 1 for all SERCAA imager sensor data.

Image file data are stored in Tagged Image File Format (TIF), therefore an alternative way to determine image dimensions is to read the TIF header and examine the width and height fields.

Image pixel values represent either counts or albedo for visible data, and brightness temperatures for thermal infrared data. Table B-1 summarizes the attributes of the SERCAA image data values.

Table B-1 Satellite image characteristics

Satellite ID (SSS)	Spectral Channel (CCC)	Channel Type	Wavelength Band	Physical Value
F10 or F11	001	Visible	0.4 - 1.1 μm	Counts[1]
	002	Long-Wave IR	10 -12 μm	Brightness Temp.[2]
N11 or N12	001	Visible	0.63 μm	Albedo[3]
	002	Near-IR	0.86 μm	Albedo
	003	Mid-Wave IR	3.7 μm	Brightness Temp.
	004	Long-Wave IR	10.7 μm	Brightness Temp.
	005	Long-Wave IR	11.8 μm	Brightness Temp.
M03	001	Visible	0.5 - 0.75 μm	Counts
	002	Long-Wave IR	10.5 - 12.5 μm	Brightness Temp.

[1] Visible counts range from 0 - 255. High counts denote highly reflective surfaces and low counts denote poorly reflective surfaces.

[2] Brightness temperatures are byte-encoded such that the range 0 - 255 corresponds to the temperature range 327.5 K to 200.0 K. The relation between byte values and temperature is linear over this range; the conversion from byte value B to brightness temperature T is given by the relation:

$$T = -0.5\,B + 327.5.$$

[3] Albedo values are byte-encoded such that the range 0 - 255 corresponds to the albedo range 0 - 100%. The relation between byte values and percent albedo is linear; the conversion from byte value B to percent albedo A is given by the relation

$$A = 0.392\,B.$$

Level 1: Latitude-Longitude File

Latitude-longitude filenames as they appear on tape have the following naming convention:

SSS_LAT_ROI_DDD_HH.Dat

where

SSS - Satellite identifier:

F10	DMSP F-10
F11	DMSP F-11
N11	NOAA-11
N12	NOAA-12
M03	METEOSAT-3

LAT - a constant that identifies the file as a latitude-longitude file

ROI - Region of Interest for which the latitude-longitude file is valid:

CNS for Central and Northern South America Area

DDD - Julian day of satellite data for which the Earth locations are valid

HH - UTC hour of the satellite data for which the Earth locations are valid

File and Record Structure

Latitude-longitude Earth location files contain fixed-length records, the number and size of which depend on both the size of the corresponding image files and the satellite type. There is always one latitude-longitude record corresponding to each satellite image file record, where a satellite image file record contains one image scan line of information. However, latitude-longitude data are subsampled, relative to the sensor data, along a scan line. There is one latitude-longitude pair for every n^{th} image pixel, where n is a function of satellite. A linear interpolation is used to compute Earth location for intermediate pixels between latitude-longitude reference points.

The information necessary for interpreting a latitude-longitude file record is contained in the Related Entries (RE) SDB information file provided with the tape, under the heading of LATLON:

LL_REC_LEN	Record length in bytes.
LL_LINE_INTERVAL	The number of image file records per lat-lon record. For the March 1994 data set this number is always 1.
LL_ELEM_INTERVAL	The subsampling rate of lat-lon information relative to the corresponding satellite data. For example, if LL_ELEM_INTERVAL = 40, there is one latitude-longitude pair for every 40th image pixel in the scan line (i.e., for pixels 1, 41, 81, ...). Linear interpolation is required to retrieve Earth location information for intermediate pixels 2-40, 42-80, ...
LL_ELEM_PER_LINE	This is the number of latitude-longitude elements per latitude-longitude file record.

A latitude-longitude file data element is a 4-byte structure that contains the scaled latitude and longitude for a given pixel. Thus the length of a latitude-longitude file record in bytes is given by:

LL_REC_LEN = 4 * LL_ELEM_PER_LINE

where the 4 bytes consist of two 16-bit integer variables: LONG and LAT. The storage convention is as follows:

LONG	Pixel longitude * 128. To obtain the floating-point longitude, FLONG = LONG / 128. Longitude range is -180^{o} to 180^{o}, positive east.
LAT	Pixel latitude * 128. to obtain floating-point latitude, FLAT = LAT / 128. Latitude range is -90^{o} to 90^{o}, positive north.

Level 1: Angles File

The angles filenames as they appear on tape have the following naming convention:

SSS_ANG_ROI_DDD_HH.Dat

where

SSS - Satellite identifier:

F10	DMSP F-10
F11	DMSP F-11
N11	NOAA-11
N12	NOAA-12
M03	METEOSAT-3

ANG - a constant that identifies the file as an angles file
ROI - Region of Interest for which the angles file is valid:
CNS for Central and Northern South America Area
DDD - Julian day of satellite data for which the angles are valid
HH - UTC hour of the satellite data for which the angles are valid

File and Record Structure

Angle files contain fixed-length records, the number and size of which depend on both the size of the corresponding image files and the satellite type. There is always one angles record corresponding to each satellite image file record, where a satellite image file record contains one image scan line of information. However, angle data are subsampled, relative to the sensor data, along a scan line. There is one set of angles for every n^{th} image pixel, where n is a function of satellite. A linear interpolation is used to compute angle values for intermediate pixels between angle reference points.

The information necessary for interpreting an angles file record is contained in the Related Entries (RE) SDB information file provided with the tape, under the heading of ANGLES:

ANG_REC_LEN	Record length in bytes.
ANG_LINE_INTERVAL	The number of image file records per angles record. This number is almost always 1.
ANG_ELEM_INTERVAL	The subsampling rate of angles information relative to the corresponding satellite image. For example, if ANG_ELEM_INTERVAL = 8, there is one set of angles valid for every eighth image pixel in the scan line (i.e., for pixels 1, 9, 17, 25, ...). Linear interpolation is required to retrieve angles information for intermediate pixels 2-8, 10-16, 18-24, ...
ANG_ELEM_PER_LINE	This is the number of angles elements per angles file record.

An angles file data element is a 12-byte structure containing three angles that define the satellite and solar viewing geometry for a given pixel. Thus the length of an angles file record in bytes is given by:

ANG_REC_LEN = 12 * ANG_ELEM_PER_LINE

where the 12 bytes consist of three 32-bit floating-point variables: SATZEN, SOLZEN, and AZIMUTH corresponding to the satellite zenith, the solar zenith, and the satellite/solar azimuth angles respectively (Figure B-1). Note: Angle files were generated on a VMS computer. To interpret these floating-point numbers on a UNIX machine it is necessary to convert from VMS to IEEE floating-point formats. Most UNIX operating systems provide a utility to perform this conversion. Angle measurement conventions are as follows:

SATZEN	Scene satellite zenith angle, 0^o - 90^o.
SOLZEN	Scene solar zenith angle, 0^o - 180^o.
AZIMUTH	Relative angle between the solar and satellite azimuth angles, 0^o - 359^o. When AZIMUTH = 0^o, the sun is directly behind the satellite (i.e., the viewed point, the satellite, and the sun are collinear). When AZIMUTH = 180^o, the satellite is looking directly into the sun (the satellite squints to compensate).

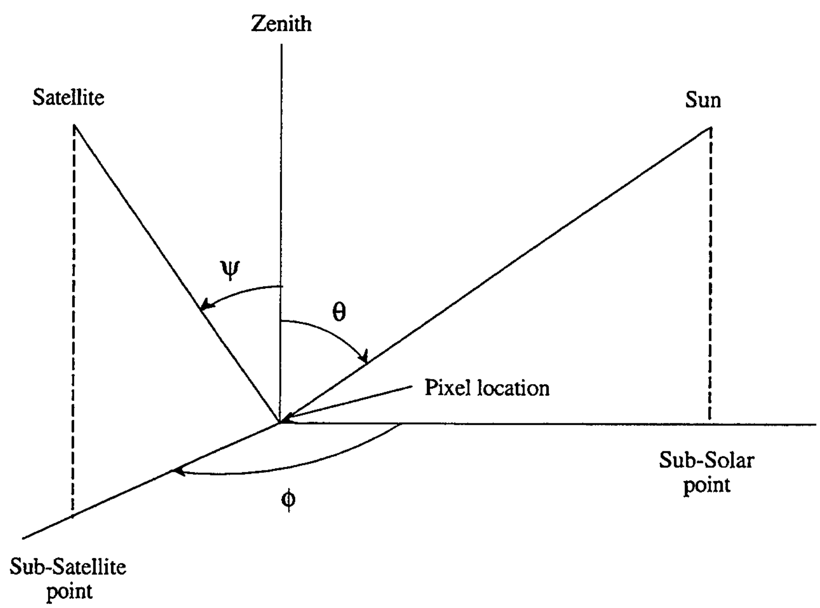


Figure B-1. Satellite-Earth-Solar Geometry (after Taylor and Stowe, 1984)

ψ - satellite zenith angle

θ - solar zenith angle

ϕ - sun-satellite azimuth angle

Level 2: **Nephanalysis Products**

Nephanalysis products are stored as bit-encoded byte values known as MCF (cloud Mask and Confidence Flag). MCF filenames as they appear on tape have the following naming convention:

SSS_MCF_ROI_DDD_HH.Dat or .Tif

where

SSS - Satellite identifier:

F10 DMSP F-10
F11 DMSP F-11
N11 NOAA-11
N12 NOAA-12
M03 METEOSAT-3

MCF - a constant that identifies the file as an MCF file
ROI - Region of Interest for which the product is valid:
CNS for Central and Northern South America Area
DDD - Julian day for which the product is valid
HH - UTC hour for which the product is valid
Dat - Raw product file format
Tif - TIF file format

File and Record Structure

Level 2 processing is performed on square arrays of image pixels, therefore the size of the resultant MCF product files is an integral number of the analysis array size. MCF files contain fixed-length records, the number and size of which depends on both the size of the corresponding image files and the satellite type. The following table specifies how to determine the record size and number of records in an MCF file. Let NCOLS and NROWS be the number of columns and rows, respectively, in the corresponding satellite image file; then:

If the image satellite id is:	Then the MCF file record size is:	And the number of lines is:
DMSP F10 or F11	NCOLS - MOD(NCOLS, 16)	NROWS - MOD(NROWS, 16)
NOAA 11 or 12	NCOLS - MOD(NCOLS, 32)	NROWS - MOD(NROWS, 32)
METEOSAT 3	See Associated RE File or TIF Header	See Associated RE File or TIF Header

where MOD is the FORTRAN modulus function (e.g., if an F10 pass has 1465 columns per scan line, then the MCF record size is 1456). The MCF file is stored in Tagged Image File Format (TIF), therefore an alternative way to determine file dimensions is to read the TIF header and examine the width and height fields.

The format of an MCF file is the same regardless of the satellite platform it was derived from. The first byte of the first record of the MCF file corresponds to the first byte of the first record in the corresponding image data file. Across each scan line there is a one-to-one correspondence between the image and MCF files out to the number of bytes computed above for each record. As can be seen in the above table, the MCF and image file sizes are not always the same. However, the two files are always aligned with respect to the upper-left corner of each.

There is one 8-bit MCF byte per analyzed image pixel. MCF bytes are bit-packed according to the following convention:

Bit 0 (least significant) is the cloud/no-cloud bit. If bit 0 is off, the corresponding image pixel is clear; if bit 0 is on, it is completely cloudy.

Bit 1 is the low cloud bit. If bit 1 is on, the pixel contains low cloud as determined by an appropriate spectral (or other) signature test.

Bit 2 is the thin cirrus cloud bit. If bit 2 is on, the pixel contains cirrus as determined by an appropriate spectral (or other) signature test.

Bit 3 is the cumulonimbus bit. If bit 3 is on, the pixel contains thunderstorm clouds.

Bit 4 is the partly cloudy bit. If bit 4 is on, the pixel is partly cloudy. If bit 4 is on, bit 0 is off. DMSP data are used exclusively to determine partly cloud conditions.

Bit 5 is the bad data bit. It is set whenever satellite data are missing or unreliable. If set, all other bits should be ignored.

Bits 6 and 7 contain the confidence level attached to the accuracy of the cloud/no-cloud decision for the corresponding cloudy image pixel. Confidence levels are rated as 0 for missing data, 1 for low confidence, 2 for mid-level confidence, and 3 for high confidence.

Low cloud, thin cirrus, and cumulonimbus conditions are always associated with completely cloudy conditions (i.e., bit 0 will always be on in the presence of one or more of these conditions). Cloud level and cloud type are not detected under partly cloudy conditions (i.e., if bit 4 is on, bits 1 through 3 will be off).

Example:

MCF byte	1	1	0	0	0	1	0	1	(C5 in hex)
bit position	7	6	5	4	3	2	1	0	

The corresponding image pixel is classified as cloud covered (bit 0) with thin cirrus (bit 2) that has been detected with a high level of confidence (bits 6 and 7).

Level 3: Layered Product

The layered product filename as it appears on tape has the following naming convention:

SAT_LYR_ROI_DDD_HH.DAT

where:

SAT - Satellite identifier:
- F10 DMSP F-10
- F11 DMSP F-11
- N11 NOAA-11
- N12 NOAA-12
- M03 METEOSAT-3

LYR is a constant that denotes the file is a layered product

ROI - Region of Interest:
- CNS for Central and Northern South America Area

DDD - Julian day

HH - GMT hour

File Structure

The layered product file contains 32265 (135 rows x 239 columns) record structures.

Record Structure

Each record contains data values valid for one grid point within a 135 (rows) by 239 (columns) two-dimensional grid. The grid is superimposed on a hemispheric secant polar stereographic map projection. Grid resolution is based on a whole mesh grid spacing of 381 km at 60 degrees latitude and nested grids are defined in terms of the number of grid cells that fit within a whole mesh grid. The layered product grid is a 1/16th mesh grid (i.e., 16 by 16 grid cells per whole mesh box.)

Table B-2 summarizes the contents of each record. Figure B-2 contains the C data structure that was used to create the data file.

Important information for reading the layering output file

The layering output file was written using an "fwrite" statement with a byte length field of "sizeof(LAYER_OUTPUT)" where LAYER_OUTPUT is the output structure contained in Figure B-2. Note that while the sum of the structure elements in Table B-2 is 55 bytes the sizeof(LAYER_OUTPUT) operator evaluates to 60 bytes. The reason for the 5 byte discrepancy is the way word alignment if performed under the UNIX operating system. The beginning of fields 6-9 and 11-14 in Table B-2 are automatically word aligned causing an additional 4 bytes to be added to the structure. The remaining byte to account for the 5 byte difference comes from rounding up the odd number of bytes in the structure. Thus, to read this file it is necessary to either use the sizeof operator in the read statement (preferred) or to hard-wire a value of 60 bytes.

Table B-2: Layered Product Record Structure

Field	Description	Units	Range	Missing or bad value	Byte length
1	Absolute 16th-mesh row number (i)		1-1024		2
2	Absolute 16th-mesh column number (j)		1-1024		2
3	SDB IR entry number			0	2
4	Julian day (ddd)				4
5	UTC (hhmm)		0-2359		2
6-9	Cloud temperature variance for each layer	GS*100			8
10	# pixels in grid box			0	2
11-14	# pixels in each layer			0	8
15-18	Cloud top temperature for each layer	GS*100			4
19-22	Cloud type for each layer		0-1		4
23-26	# low cloud pixels in each layer				4
27-30	# thin cirrus pixels in each layer				4
31-34	# precipitating-cloud pixels in each layer				4
35	Sunrise time		0-235		1
36	Sunset time		0-235		1
37	Satellite platform ID				1
38	# data dropouts in grid box				1
39	# partially cloud-filled pixels				1

```
/* Layering output structure

   Daniel Peduzzi (AER) 9/27/94
     structure content by Robert P. d'Entremont (AER) 9/1994
*/

#ifndef NCLASSES
#  define NCLASSES (4)
#endif

#ifndef _LAYER_OUTPUT
#define _LAYER_OUTPUT

#define BYTE unsigned char

typedef  struct {

  short i;                          /* 16th-mesh absolute row (1-1024)                 */
  short j;                          /* 16th-mesh absolute column (1-1024)              */

  short sdb_ir_entry;               /* SDB entry number corresponding to IR data       */
  int yyddd;                        /* Sensor data Julian day                          */
  short hhmm;                       /* Sensor data valid time (UTC) hhmm               */

  short layer_var[NCLASSES];        /* Temperature variance*100 for cloud layer i      */
  short num_pixels;                 /* Total # of pixels in 16th-mesh box              */
  short n_layer_pix[NCLASSES];      /* Total # pixels in layer i                       */
  BYTE meantemp[NCLASSES];          /* Mean cloud top temperature for layer i          */
  BYTE cloud_type[NCLASSES];        /* Cloud type for layer i (1 or 2)                 */
  BYTE low_cloud[NCLASSES];         /* # low cloud pixels in layer i                   */
  BYTE thin_cirrus[NCLASSES];       /* # thin cirrus pixels in layer i                 */
  BYTE precip[NCLASSES];            /* # precipitating-cloud pixels in layer i         */

  BYTE sunrise;                     /* Sunrise time (UTC) (0-235)                      */
  BYTE sunset;                      /* Sunset time (UTC) (0-235)                       */
  BYTE vid;                         /* Satellite vehicle (platform) ID                 */
  BYTE dropouts;                    /* Total # of data dropouts in 16th-mesh box       */
  BYTE partial;                     /* Total # of partially-cloud-filled pixels        */

} LAYER_OUTPUT;

#undef BYTE

#endif
```

Figure B-2: Level 3 data structure

Level 4: Integrated Product

The integrated product filename as it appears on tape has the following naming convention:

ALL_IAN_ROI_DDD_HH.Dat

where

ALL and IAN are constants (Integrated ANalysis from ALL sensors)
ROI - Region of Interest for which the product is valid
    Possible values:
        CNS for the Central and Northern South America Area
DDD - Julian day for which the integrated product is valid
HH - GMT hour for which the integrated product is valid

File Structure

The integrated product file contains 32,265 records (239 columns by 135 rows), each 64 bytes in length.

Record Structure

Each record contains data values valid for one grid point within a 135 (rows) X 239 (columns) 2-D grid. The grid is superimposed on a hemispheric secant polar stereographic map projection. Grid resolution is based on a whole-mesh grid spacing of 381 km at 60^{o} latitude and nested grids are defined in terms of the number of grid cells that fit within a whole mesh grid. The integrated product grid is a $1/16^{th}$ mesh grid (i.e., 16 X 16 cells per whole mesh box).

Table B-3 summarizes the contents of each record. All values are 16-bit integers and one grid cell occupies 64 bytes. Figure B-3 contains the C data structure used to create the output file.

Table B-3. Integrated Product Record Structure*

Field	Description	Units	Range	Missing or bad value	Comments
1	Absolute 16th-mesh column number (i)		227 - 451		
2	Absolute 16th-mesh row number (j)		13 - 395		
3	Number of cloud layers in (i,j)		0 - 4	-999	
4	Total cloud fraction for (i,j)	Percent	0 - 100	-999	
5-8	Cloud fraction by layer for (i,j)	Percent	0 - 100	-999	
9-12	Cloud top temperature by layer	K*10	2000-3275	-999	
13-16	Cloud top height by layer	Meters	0-13500	-999	
17-20	Cloud type by layer		0 - 9	-999	See Table B-4
21	Total cloud fraction error for (i,j)	Percent	0 - 100	-999	
22-25	Layer cloud fraction error for (i,j)	Percent	0 - 100	-999	
26-29	Layer confidence flags for (i,j)	Flag*10	10 - 30	-999	Discrete values for low to high confidence
30-32	Database entry numbers for input satellite analyses				Corresponds to directory names on tar tape

*all values are 16-bit integers

Table B-4. Cloud Type Codes

Cloud Type Code	Cloud Type
0	No Cloud
1	Cirrus
2	Cirrostratus
3	Altocumulus
4	Altostratus
5	Stratocumulus
6	Stratus
7	Cumulus
8	Cumulonimbus
9	Nimbostratus

```
/* CNSA definitions */

#define NLINE 135
#define NCOL 239
#define NLYRS 4


#define MIN_I 413
#define MIN_J 877

typedef unsigned char byte;

/* integration output structure */

typedef struct {
   short i;                        /* absolute 16th mesh coord         */
   short j;
   short nlayers;                  /* number of layers                 */
   short fraction;                 /* total cloud fraction             */

   short lyr_frc[NLYRS];           /* layer cloud fraction             */
   short t_cld[NLYRS];             /* layer cloud top temp (K*10)      */
   short z_cld[NLYRS];             /* layer cloud top height (m)       */
   short cld_typ[NLYRS];           /* layer cloud type                 */

   short error;                    /* total cloud amount error         */
   short lyr_err[NLYRS];           /* layer cloud amount error         */
   short conf[NLYRS];              /* layer confidence measure         */
   short sdb_entry[3];             /* input entry number(s)            */
} INTEGRATION;
```

Figure B-3: Integration output data structure

Appendix C

Data Extraction Guide

****SERCAA DATA SET RELEASE TO DNA****

****

What should I have ?

DNA_RELEASE.TXT	This document.
(2) 8 mm D8-112 tapes	One tape, labeled DNA MAR94 CNS IA, contains the SERCAA Integrated Analysis (SIA) data files. The other tape, labeled DNA MAR94 CNS ENTRIES contains the Related Entry (RE) data (which consists of Satellite, Latitude/Longitude, Angles(Geometry) and Product(cloud mask) data files.

****

What type of tape drive was used ?

A SUN Exabyte EXB-8500 8 mm tape drive recording in high density mode (5 gig).

****

What utility was used to create the release tapes ?

The data were placed on the tapes using a SUN SPARC II running SUN OS 4.1.2. The following tar command syntax was used:

```
sun% tar cvBf /dev/nrst8 somedirectory
```

****

How are the data arranged on the release tape ?

The data on the SIA tape are contained in 10 tar files. Each of these tar files represents a directory that contains all the SIA data for a particular day (day 94079 through day 94091). Each directory name follows the convention:

CYYJJJ

where:

C = century (9 for 19XX)
YY = year
JJJ = Julian day

A SIA file and SIA SDB information file exists for each hour that an analysis was performed. Each SIA file has been named using the following convention:

Positions 1-4 Platform:

all_ = All satellite platforms are used to create a SIA.

Positions 5-8 Type of file:

ian_ = integrated analysis file
sdb_ = SERCAA data base (SDB) information file

Positions 9-12 Region of interest:

(Given in 16th-mesh coordinates)
eas_ = East Asia Area (EASA). (i,j) = (227,13) to (451,395)
can_ = Canada Area (CANA). (409,597) to (557,711)
cns_ = Central, Northern South America Area (CNSA). (413,877) to (651,1011)
emd_ = Eastern Mediterranean, Desert Area (EMDA). (731,353) to (863,505)

Positions 13-16 Julian day:

081_ = Julian day 081 etc. ...

Positions 17-18 Hour:

00 = SIA for hour 00 etc. ...

Positions 19-22 Extension:

.dat = raw-format file extension

Example: all_ian_eas_081_10.dat

The RE tape contains about 379 tar files. Each of these tar files represent a directory that contains all the related data used as input to create at least one of the SIA data files. Each directory name follows the convention:

ENTRY/

where:
ENTRY = the SDB entry number

Each RE file has been named following these guidelines:

Positions 1-4 Platform:

n11_ = NOAA N_11
n12_ = NOAA N_12
f10_ = DMSP F_10
f11_ = DMSP F_11
g04_ = GMS-4
m03_ = METEOSAT-3

Positions 5-8 Type of file:

001_ = satellite data channel 1
002_ = satellite data channel 2
...
...
005_ = satellite data channel 5
lat_ = latlon data
ang_ = angles data
mcf_ = cloud mask data
sdb_ = SDB information file

Positions 9-12 Area of data:

eas_ = East Asia Area (EASA)
can_ = Canada Area (CANA)

cns_ = Central and Northern South America Area (CNSA)
emd_ = Easter Mediterranean, Desert Area (EMDA)

Positions 13-16 Julian day:

081_ = Julian day 081 etc. ...

Positions 17-18 Hour:

00 = hour of the data

Positions 19-22 Extension:

.dat = raw data
.tif = tif formatted data

Examples:

f10_001_eas_150_14.tif
f10_002_eas_150_14.tif
f10_lat_eas_150_14.tif
f10_ang_eas_150_14.tif
f10_mcf_eas_150_14.tif
f10_sdb_eas_150_14.tif

Refer to separate listing sheet labeled MAR94.IA.TAR.LIST for a listing of the IA tape contents. Run the provided script, "list_tar", to generate a listing of the RE tape.

****

What are related data items ?
What is the SDB entry number ?
What are related entries ?

The SDB registration process is a process that automatically places descriptive data items about a satellite scan into the SDB. The SDB registration process allocates a group of unique entry numbers to be used as place holders for all of the related data items for a given satellite scan. The related data items consists of satellite, latitude/longitude, angles (Geometry) and product(cloud mask) data. As an example, if a DMSP F_11 scan was to be registered in the SDB, the registration process would request for a group of five contiguous entry numbers(i.e. 1001-1005). These five entry numbers would be used as place holder for the following related data items:

1001	f11 visible channel
1002	f11 infrared channel
1003	latitude/longitude data
1004	angles(geometry) data
1005	product data

The "SDB entry number" is the first entry number of the group of entry numbers provided by the registration process. The first entry number is used to "key" into the related data items for that group. In the example provided above the SDB entry number would be 1001.

The release process uses the SDB entry number in each group to logically divide the data into separate directories (i.e. the directory name is first SDB entry number for each group of entry numbers). Using the example provided above the directory named "1001/" contains all the related data items for that group (i.e. the directory contains the data for entry 1001 through entry 1005).

To build a SIA it is necessary to use as input, related data items from one or more satellite scans and/or satellite platforms. The SDB entry number is used to keep track of all inputs to the SIA. The list of related entries are given as SDB entry numbers.

****

How do I get a particular SIA data set ?

You must use the UNIX tar utility to extract the data from the tape. By using the tar utility you may extract individual files or the entire directory. Use the MAR94.IA.TAR.LIST to determine where to position the tape and then use the appropriate tar command to extract the files you want. For example, if you want to extract all of the SIA data files from the first and second tar files, the following commands could be used:

```
% pwd
/users/smith
% mkdir data
% cd data
% tar xvf /dev/rst8 994081
% tar xvf /dev/rst8 994082
```

Upon completion all of the SIA data for day 081 would reside in directory /users/smith/data/994081 and all the SIA data for day 082 would reside in directory /users/smith/data/994082.

****

What is the SDB information file ?

The SDB information file is a text file containing selected SDB record items that help describe the actual data. The SIA SDB information file shows what data went into creating the SIA by listing the related entries. The RE SDB information file lists information about the satellite images, the latlon file, the angles file and the product file(s).

The following is an example SIA SDB information file:

```
[IA]
ZULU_YYJJJ:=94081                        : Year, Julian day of SIA
ZULU_HH:=10                              : Hour of SIA
ROI:=CNS                                 : Region of Interest
NUM_RELATED_LAYER:=3                     : Number for related entries
RELATED_LAYER_1:= 4148                   : 1st related SDB entry number
RELATED_LAYER_2:= 7199                   : 2nd related SDB entry number
RELATED_LAYER_3:= 8988                   : 3d related SDB entry number
TDISK:=SDB_Int:
TDIR:=[SERCAA.DATA.994081]
FILE_IA_1:=ALL_IAN_CNS_081_08.Dat        : SIA file name
SDB_SET:=MAR94                           : Set identifier March of 1994
```

The following is an example RE SDB information file:

```
[SATIMG]
SAT_CODE:=16                             : Satellite code
ZULU_YYJJJ:=94081                        : Year, Julian day of scan
ZULU_HHMMSS:=82252                       : Time of scan
NUM_LINES:=1375                          : Number of lines
```

```
ELEM_PER_LINE:=409                                                  : Elements per line
BYTES_PER_ELEM:=1                                                   : Bytes per element
7199:=AVH$005:[SERCAA.DATA.994081]N11_001_CNS_081_08.TIF            : Channel 1 file
7200:=AVH$005:[SERCAA.DATA.994081]N11_002_CNS_081_08.TIF            : Channel 2 file
7201:=AVH$005:[SERCAA.DATA.994081]N11_003_CNS_081_08.TIF            : Channel 3 file
7202:=AVH$005:[SERCAA.DATA.994081]N11_004_CNS_081_08.TIF            : Channel 4 file
7203:=AVH$005:[SERCAA.DATA.994081]N11_005_CNS_081_08.TIF            : Channel 5 file

[LATLON]
LL_REC_LEN:=204                                                     : Record length in bytes
LL_LINE_INTERVAL:=1                                                 : Sub-sample line interval
LL_ELEM_INTERVAL:=8                                                 : Sub-sample element interval
LL_ELEM_PER_LINE:=51                                                : Latlon pairs per line
LL_FILE:=AVH$005:[SERCAA.DATA.994081]N11_LAT_CNS_081_08.DAT         : latitude/longitude file

[ANGLES]
ANG_REC_LEN:=612                                                    : Record length in bytes
ANG_LINE_INTERVAL:=1                                                : Sub-sample line interval
ANG_ELEM_INTERVAL:=8                                                : Sub-sample element interval
ANG_ELEM_PER_LINE:=51                                               : Angles triplets per line
ANG_FILE:=AVH$005:[SERCAA.DATA.994081]N11_ANG_CNS_081_08.DAT        : Angles file

[PRODUCT]
7206001:=sdb$prd:[SERCAA.DATA.994081]N11_MCF_SET_081_08.TIF         : Cloud mask file
```

****

How do I know which RE data went into a particular SIA ?

There are two ways to determine which RE data sets went into a particular SIA. The first way is reference the SIA SDB information file. Each "RELATED_LAYER" listed is a reference, by SDB entry number, to the RE data. Use the referred SDB entry number to retrieve the related data from the RE data tape.

For example, refer to the above SIA SDB information file. The "RELATED_LAYERED_1:=4148" line implies that SDB entry number 4148 and the related data items for entry 4148 (along with SDB entry numbers 7199 and 8988) were used to create "ALL_IAN_CNS_081_08.Dat".

The second way is to read the header information from the SIA file (Please refer to the DATA_DESCRIPTION).

****

How do I get the RE data files ?

Once you have examined the SIA SDB information file and you have identified the related entry numbers, you must use the UNIX tar utility to extract the data from the tape. By using the tar utility you may extract individual files or the entire directory. Use a tape contents list generated using the "list_tar" script to determine where to position the tape and then use the appropriate tar command to extract the files you want. For example, if you want to extract all of the RE data files from the first tar file, the following commands could be used:

```
% pwd
/users/smith
% mkdir data
% cd data
% tar xvf /dev/rst8 7199
```

Upon completion of this command all of the RE data related to SDB entry number 7199 would reside in directory /users/smith/data/7199.

****

For the following question please refer to the example SDB information files as needed.

****

What is the format of the satellite data and how do I access it?

The dimensions of the satellite data are defined by the three parameters, NUM_OF_LINES, ELEM_PER_LINE and BYTES_PER_ELEM . To access the data use the following logic.

If the file extension is ".dat"
then use the appropriate C or FORTRAN read statements.

If the file extension is ".tif"
then use a TIF reader or TIF library (you may view the
images by using the public domain application, XV).

For a detailed explanation, refer to Appendix B.

****

What is the format of the latlon data and how do I access it?

The latlon data are sub-sampled. The dimensions are defined LL_LINE_INTERVAL, LL_ELEM_INTERVAL and LL_ELEM_PER_LINE. LL_ELEM_PER_LINE defines the number of longitude/latitude pairs per line. Each pair is four bytes (two bytes lon, two bytes lat). To access the data use the appropriate C or FORTRAN read statements.

For a detailed explanation, refer to Appendix B.

****

What is the format of the angles data and how do I access it?

The angles data are sub-sampled. The dimensions are defined by ANG_LINE_INTERVAL, ANG_ELEM_INTERVAL and ANG_ELEM_PER_LINE. ANG_ELEM_PER_LINE
defines the number of triplets (satellite-zenith/solar-zenith/azimuth) per line. Each item in the triplet is a float data type. To access the data use the appropriate C or FORTRAN read statements.

For a detailed explanation, refer to Appendix B.

P601

Data Save Documentation Report No. 5

ADVANCED GEOPHYSICAL ENVIRONMENT SIMULATION TECHNIQUES

Task 1: Satellite Data Sets for Worldwide Cloud Prediction

This data documentation report covers data set generation
for the DNA region of interest:

Eastern Mediterranean Desert Area (EMDA)

for the period:

12-21 March 1994

Contract Number F19628-94-C-0106

issued by:

Electronic Systems Center
Air Force Materiel Command
Hanscom AFB, MA 01731

Submitted by:

Atmospheric and Environmental Research, Inc.
840 Memorial Drive
Cambridge, MA 02139

25 May 1995

David B. Hogan
Gary B. Gustafson
Principal Investigators

1.0 Introduction

This Data Documentation Report provides a description of the fifth data save made in accordance with the revised statement of work for Satellite Data Sets for Worldwide Cloud Prediction Models. It is intended to provide a description of the data set, its format, how the data were collected and processed, and the algorithms used to generate it. The data set consists of raw satellite data and analyzed products produced by the SERCAA cloud analysis algorithms. The period covered is 12-21 March 1994 for the DNA region of interest: Eastern Mediterranean Desert America (EMDA). This region covers the following (i,j) 16th mesh grid coordinates: 731,353 - 863,505. All available data from those dates are included. These data were processed specifically for DNA using software developed from the SERCAA cloud analysis algorithms described by Gustafson et al (1994). Substantial modifications were required to the Cloud Layering and Analysis Integration modules to accommodate the high volume of data included in this data set. Two tapes are provided, one with Level 1, 2 and 3 products and the second with Level 4.

2.0 Processing Environment

Satellite data processing for this data set used the SERCAA cloud analysis algorithms described by Gustafson et al. (1994). Multisource data from the DMSP F10 and F11, NOAA-11 and NOAA-12, and METEOSAT-4 satellites were used. Data sources were as follows: DMSP - National Geophysical Data Center (NGDC), Boulder, CO; NOAA - National Climatic Data Center (NCDC), Ashville, NC; METEOSAT - Phillips Laboratory direct readout ground station. All polar-orbiting satellite data were obtained by the Phillips Laboratory and were received on tape in various formats. All data processing was performed on the Air Force Interactive Meteorological System (AIMS) at the Phillips Laboratory. The SERCAA cloud analysis algorithms use four levels of data processing as summarized in Figure 1.

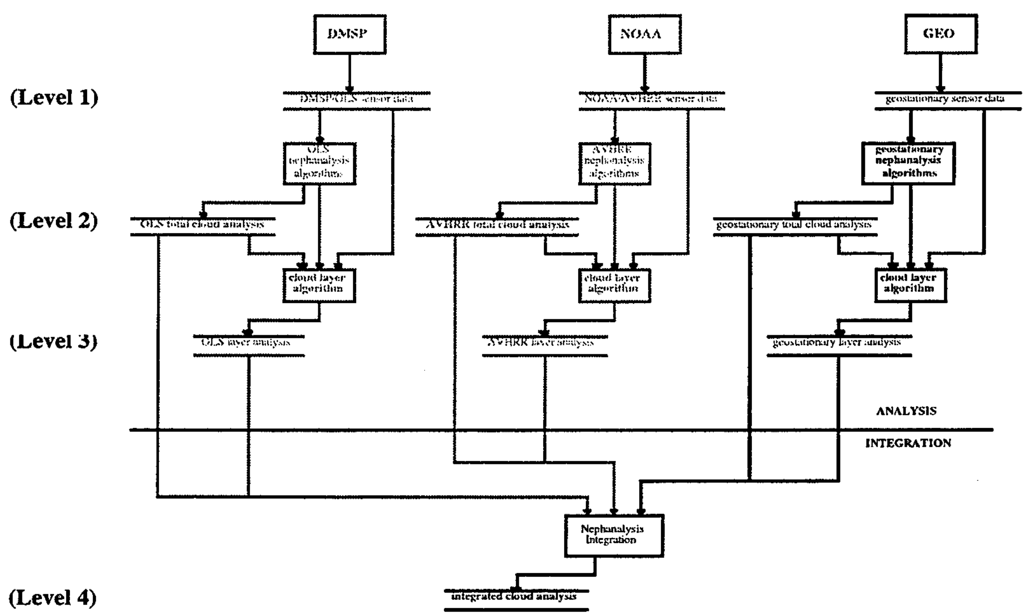


Figure 1. SERCAA data flow and processing levels

Level 1 processing consists of data ingest. Tape data are processed through separate data source and format dependent ingest programs. All data are then stored in a standard format as flat files in the original satellite scan projection. The number of elements and rows correspond to the number of pixels in a scan line and the number of scan lines respectively. Data are maintained on AIMS through customized SERCAA Database (SDB) management software. Level 1 data products consist of separate files for each sensor channel plus two additional files containing Earth location and satellite/solar geometry information. One set of files is created for each polar satellite pass and geostationary satellite scan. Satellite data characteristics are summarized in Table 1. In cases where visible and infrared channel resolution differ, the higher resolution data are subsampled to match the coarser resolution data (e.g., METEOSAT visible data are subsampled by a factor of two to match the IR data resolution). Earth location data consist of latitude-longitude pairs maintained at a subsampled resolution relative to the sensor data. For each sensor scan line, one latitude-longitude pair is provided for every n^{th} pixel, where n varies with satellite. Geometry information are also subsampled in the same ratio as the Earth location information and consist of three angles: satellite zenith, solar zenith, and sun-satellite azimuth. Ingest products are described more completely in Section 2 of Gustafson et al. (1994).

Table 1. Sensor Channel Data Attributes During SERCAA

Satellite	Sensor	Channel (µm)	Data Format	Resolution[1] (km)	Bits per Pixel[2]	Pixels per Scan Line
DMSP	OLS	0.40-1.10	counts	2.7	6	1464
		10.5-12.6	EBBT	2.7	8	1464
NOAA	AVHRR	0.58-0.68	percent albedo	4.0	10	409
		0.72-1.10	percent albedo	4.0	10	409
		3.55-3.93	EBBT	4.0	10	409
		10.3-11.3	EBBT	4.0	10	409
		11.5-12.5	EBBT	4.0	10	409
METEOSAT	VISSR	0.55-0.75	counts	2.5	8	5000
		10.5-12.6	EBBT	5.0	8	2500

[1]Sensor resolution at satellite subpoint that will provide global coverage.

[2]AVHRR radiance data are transmitted at 10-bit resolution, however, the SERCAA development system could only accommodate 8-bit brightness temperature data (although the full 10-bit resolution is used in the radiance-to-brightness-temperature transformation).

Level 2 processing consists of sensor-specific nephanalysis algorithms. Level 1 sensor data from DMSP, NOAA, and the METEOSAT satellites are processed through separate algorithms as indicated in Figure 1. Each Level 1 data set is analyzed using the appropriate nephanalysis algorithm and results placed in a Level 2 output file and tagged with the sensor data valid time. One output file is generated for each nephanalysis run and nephanalysis results are stored in the original satellite scan projection with one byte of information for each pixel. Each byte is bit-packed according to the structure in Table 2. Thus for each set of Level 1 products generated from a satellite pass, one Level 2 product file is generated.

Table 2. Cloud Analysis Algorithm MCF File Bit Assignments

Bit	Assignment	Description
0	Cloud Mask	ON = Cloud-Filled OFF = Cloud-Free
1	Low Cloud	ON = Low Cloud Found
2	Thin Cirrus Cloud	ON = Thin Cirrus Cloud Found
3	Precipitating Cloud	ON = Precipitating Cloud Found
4	Partial Cloud	Only used by DMSP algorithm
5	Data Dropout	ON = Missing or Unreliable Data
6	Confidence	0 = Missing Data; 1 = Low;
7	Flag	2 = Middle; 3 = High

Level 3 processing uses Level 1 and 2 products as input to stratify the cloudy regions into vertical cloud layers and to classify different cloud types. For this data set **no cloud type information is computed** due to processing time constraints. Level 3 products are remapped from the individual satellite projections to the AFGWC standard polar stereographic map projection (Hoke et al., 1981) at 16^{th} mesh grid resolution. The EMDA region of interest processed for the March 1994 data set have the following (i,j) 16^{th} mesh grid coordinates: $731 \leq i \leq 863$, $353 \leq j \leq 505$. Level 3 products are generated for each 16^{th} mesh grid cell and contain the information in Table 3. A maximum of four cloud layers can be identified for each grid cell. One Level 3 file is created for each set of Level 1 and 2 products. All Level 1, 2, and 3 products associated with a single satellite pass are related through SDB and tagged with the valid time of the Level 1 sensor data. Note that for the EMDA region, all Level 3 files are a fixed size of 133x153 grid cells.

Table 3. Cloud Typing and Layering Output

Parameter	Description
i	16^{th} mesh i coordinate for Grid Cell
j	16^{th} mesh j coordinate for Grid Cell
sdb_ir_entry	SDB entry number of input IR sensor data
ddd	Sensor data Julian date
hhmm	Sensor data valid time (UTC)
layer_var(4)	Cloud top IR variance of pixels in each layer
num_pixels	Total number of satellite pixels in 16^{th} mesh grid cell
n_layer_pix(4)	Total number of pixels in each layer
meantemp(4)	Cloud top mean IR Temperature of pixels in each layer
cloud_type(4)	Cloud type of each layer*
mean_conf(4)	Mean confidence level for each layer
low_cloud(4)	Number of low cloud pixels in this layer detected by cloud analysis algorithm
thin_cirrus(4)	Number of thin cirrus pixels in this layer detected by cloud analysis algorithm
precip(4)	Number of precipitating cloud pixels in this layer detected by cloud analysis algorithm
sunrise	Local sunrise time (UTC)
sunset	Local sunset time (UTC)
vid	Satellite vehicle (platform) ID
dropouts	Number of bad data pixels in 16^{th} mesh grid cell
partial	Number of partial cloud pixels detected by DMSP cloud analysis algorithm

*cloud type information not provided.

Level 4 processing is a clock driven process with one new Level 4 integrated analysis performed each hour. Thus, integration is differentiated from the Level 1, 2, and 3 products that are event-driven (i.e., resulting from the ingest of a new satellite pass). The integration module operates on the most recent Level 3 gridded products available from each satellite source (i.e., NOAA, DMSP, METEOSAT). Like Level 3 products, the Level 4 output files conform to the AFGWC 16th mesh grid structure; output parameters for each grid cell are summarized in Table 4.

Table 4. Analysis Integration Processed Parameters

Parameter	Description
i	16th mesh i (column) coordinate
j	16th mesh j (row) coordinate
nlay	Number of Cloud Layers
cftot	Total Cloud Fraction
cf(4)	Layer Cloud Fraction
ctt(4)	Layer Cloud Top IR Temperature (K)
ctz(4)	Layer clout top height (m)
ity(4)	Layer Cloud Type
ecft	Estimated Error in Total Cloud Fraction
ecf(4)	Estimated Error in Layer Cloud Fraction
icf(4)	Analysis Confidence Flag Index For Each Layer
sdb(3)	SDB entry number of input analyses (NOAA, DMSP, METEOSAT)

3.0 Tape Format

All data for the March 1994 EMDA data save are contained on two 8 mm tapes written in UNIX tar format. The first tape, labeled: DNA MAR94 EMD IA/RE 071-078, contains all Level 1-4 products for the first 8 days of the ten-day period. The second tape, labeled: DNA MAR94 EMD IA/RE 079-080, contains all the Level 1-4 products for the last 2 days of the ten-day period. The size of the combined Level 1, 2 and 3 products is approximately 1.8 Gbytes and the Level 4 products occupy 545 Mbytes. A UNIX script is provided to generate a listing of the contents of the tar tapes at the user's site. It may be useful to place the listing file generated by the script into an edit program to scan and search it quickly. The listings are required to locate specific data sets on the tapes.

Level 1-3 products are generated for each new pass of satellite data received during the period of the data save. Appendix A contains a chronological list of each satellite pass used to produce the March 1994 data sets. All available data for the period covered were included; any gaps in the data list are due to either missing or bad data. DMSP data quality was improved over the previously analyzed 1993 data sets and consistent with the quality of the March 1994 CNSA data. Although a few orbits of DMSP had to be dropped from the processing stream due to excessive bad or missing lines, there were no instances of the periodic drop-outs that had been ubiquitous in the 1993 data sets.

For each set of Level 1-3 products and each Level 4 file there is also an SDB Information File. These files contain descriptive metadata information extracted from the

SERCAA Database that describe the relevant attributes of the SERCAA product files. For example, information files list the number of pixels per scan line and the number of scan lines in the file. Information on subsampling ratios for the Earth location and angles files is also contained there.

Detailed descriptions of the file formats used for each output level, and the associated information files, provided for the March 1994 save (Level 1, 2, 3, and 4) are provided in Appendix B. Appendix C provides a guide for extracting data sets from tape.

4.0 References

Gustafson, G.B., R.G. Isaacs, R.P. d'Entremont, J.M. Sparrow, T.M. Hamill, C. Grassotti, D.W. Johnson, C.P. Sarkisian, D.C. Peduzzi, B.T. Pearson, V.D. Jakabhazy, J.S. Belfiore, and A.S. Lisa, 1994: Support of Environmental Requirements for Cloud Analysis and Archive (SERCAA): algorithm descriptions. PL-TR-94-2114, Phillips Laboratory, Hanscom AFB, MA, ADA283240.

Hoke, J.E., J.L. Hayes, L.G. Renninger, 1981: Map projections and grid systems for meteorological applications. AFGWC-TN-79-003, Air Weather Service, Scott, AFB, IL.

Appendix A

Chronological List of Input Satellite Data

ENTRY	DATE	TIME	ROI	SATELLITE	ELES	LINES	RESLN
6607	94070	152400	EMD	DMSP_F11	1465	1045	2.70
6612	94070	170600	EMD	DMSP_F11	1465	964	2.70
6587	94070	190400	EMD	DMSP_F10	1465	1102	2.70
6592	94070	214400	EMD	DMSP_F10	1465	951	2.70
2915	94071	013000	EMD	MET_4	640	512	5.00
5001	94071	023000	EMD	MET_4	640	512	5.00
5126	94071	033000	EMD	MET_4	640	512	5.00
5151	94071	043000	EMD	MET_4	640	512	5.00
6617	94071	045200	EMD	DMSP_F11	1465	988	2.70
5442	94071	053000	EMD	MET_4	640	512	5.00
5447	94071	063000	EMD	MET_4	640	512	5.00
6597	94071	072100	EMD	DMSP_F10	1465	898	2.70
5452	94071	073000	EMD	MET_4	640	512	5.00
5457	94071	083000	EMD	MET_4	640	512	5.00
6602	94071	090100	EMD	DMSP_F10	1465	871	2.70
5462	94071	093000	EMD	MET_4	640	512	5.00
5467	94071	103000	EMD	MET_4	640	512	5.00
5472	94071	113000	EMD	MET_4	640	512	5.00
5477	94071	123000	EMD	MET_4	640	512	5.00
5482	94071	133000	EMD	MET_4	640	512	5.00
5487	94071	143000	EMD	MET_4	640	512	5.00
6622	94071	151100	EMD	DMSP_F11	1465	968	2.70
5492	94071	153000	EMD	MET_4	640	512	5.00
5497	94071	163000	EMD	MET_4	640	512	5.00
5502	94071	173000	EMD	MET_4	640	512	5.00
5507	94071	183000	EMD	MET_4	640	512	5.00
5511	94071	193000	EMD	MET_4	640	512	5.00
6632	94071	193200	EMD	DMSP_F10	1465	1084	2.70
6637	94071	201100	EMD	DMSP_F10	1465	1124	2.70
5516	94071	203000	EMD	MET_4	640	512	5.00
5520	94071	213000	EMD	MET_4	640	512	5.00
5524	94071	223000	EMD	MET_4	640	512	5.00
5528	94071	233000	EMD	MET_4	640	512	5.00
5532	94072	013000	EMD	MET_4	640	512	5.00
3674	94072	015244	EMD	NOAA_11	409	830	4.00
5536	94072	023000	EMD	MET_4	640	512	5.00
5540	94072	033000	EMD	MET_4	640	512	5.00
3690	94072	033344	EMD	NOAA_11	409	765	4.00
5544	94072	043000	EMD	MET_4	640	512	5.00
6657	94072	043900	EMD	DMSP_F11	1465	1020	2.70
4010	94072	045137	EMD	NOAA_12	409	823	4.00
3682	94072	051414	EMD	NOAA_11	409	244	4.00
6662	94072	052000	EMD	DMSP_F11	1465	896	2.70
5549	94072	053000	EMD	MET_4	640	512	5.00
5554	94072	063000	EMD	MET_4	640	512	5.00
4018	94072	063200	EMD	NOAA_12	409	758	4.00
5559	94072	073000	EMD	MET_4	640	512	5.00
6642	94072	075000	EMD	DMSP_F10	1465	900	2.70
5564	94072	080000	EMD	MET_4	640	512	5.00

ENTRY	DATE	TIME	ROI	SATELLITE	ELES	LINES	RESLN
6647	94072	093000	EMD	DMSP_F10	1465	906	2.70
5569	94072	093000	EMD	MET_4	640	512	5.00
5574	94072	103000	EMD	MET_4	640	512	5.00
5579	94072	113000	EMD	MET_4	640	512	5.00
5584	94072	123000	EMD	MET_4	640	512	5.00
3698	94072	131139	EMD	NOAA_11	409	959	4.00
5589	94072	133000	EMD	MET_4	640	512	5.00
5594	94072	143000	EMD	MET_4	640	512	5.00
3706	94072	145223	EMD	NOAA_11	409	1012	4.00
5599	94072	153000	EMD	MET_4	640	512	5.00
6667	94072	155900	EMD	DMSP_F11	1465	1064	2.70
4026	94072	160607	EMD	NOAA_12	409	937	4.00
5604	94072	163000	EMD	MET_4	640	512	5.00
5609	94072	173000	EMD	MET_4	640	512	5.00
6672	94072	173900	EMD	DMSP_F11	1465	989	2.70
4034	94072	174617	EMD	NOAA_12	409	997	4.00
5614	94072	183000	EMD	MET_4	640	512	5.00
5618	94072	193000	EMD	MET_4	640	512	5.00
5623	94072	203000	EMD	MET_4	640	512	5.00
5627	94072	213000	EMD	MET_4	640	512	5.00
5631	94072	223000	EMD	MET_4	640	512	5.00
5635	94072	233000	EMD	MET_4	640	512	5.00
5639	94073	013000	EMD	MET_4	640	512	5.00
3714	94073	014036	EMD	NOAA_11	409	806	4.00
5643	94073	023000	EMD	MET_4	640	512	5.00
3730	94073	032123	EMD	NOAA_11	409	753	4.00
5647	94073	033000	EMD	MET_4	640	512	5.00
5651	94073	043000	EMD	MET_4	640	512	5.00
4042	94073	043016	EMD	NOAA_12	409	786	4.00
3722	94073	050154	EMD	NOAA_11	409	485	4.00
5656	94073	053000	EMD	MET_4	640	512	5.00
4050	94073	061017	EMD	NOAA_12	409	790	4.00
5661	94073	063000	EMD	MET_4	640	512	5.00
5666	94073	073000	EMD	MET_4	640	512	5.00
4058	94073	075017	EMD	NOAA_12	409	463	4.00
5671	94073	083000	EMD	MET_4	640	512	5.00
6757	94073	085800	EMD	DMSP_F10	1465	932	2.70
5676	94073	093000	EMD	MET_4	640	512	5.00
5681	94073	103000	EMD	MET_4	640	512	5.00
6762	94073	103700	EMD	DMSP_F10	1465	903	2.70
5686	94073	113000	EMD	MET_4	640	512	5.00
5701	94073	123000	EMD	MET_4	640	512	5.00

ENTRY	DATE	TIME	ROI	SATELLITE	ELES	LINES	RESLN
3738	94073	130020	EMD	NOAA_11	409	798	4.00
5696	94073	133000	EMD	MET_4	640	512	5.00
5691	94073	143000	EMD	MET_4	640	512	5.00
3746	94073	144002	EMD	NOAA_11	409	1023	4.00
5706	94073	153000	EMD	MET_4	640	512	5.00
4066	94073	154609	EMD	NOAA_12	409	656	4.00
3754	94073	162320	EMD	NOAA_11	409	165	4.00
5711	94073	163000	EMD	MET_4	640	512	5.00
4074	94073	172432	EMD	NOAA_12	409	1016	4.00
5716	94073	173000	EMD	MET_4	640	512	5.00
5721	94073	183000	EMD	MET_4	640	512	5.00
4082	94073	190702	EMD	NOAA_12	409	607	4.00
6767	94073	190800	EMD	DMSP_F10	1465	1096	2.70
5726	94073	203000	EMD	MET_4	640	512	5.00
6772	94073	214900	EMD	DMSP_F10	1465	943	2.70
5730	94073	223000	EMD	MET_4	640	512	5.00
3762	94074	012830	EMD	NOAA_11	409	775	4.00
5734	94074	023000	EMD	MET_4	640	512	5.00
3770	94074	030900	EMD	NOAA_11	409	814	4.00
5738	94074	033000	EMD	MET_4	640	512	5.00
4090	94074	040903	EMD	NOAA_12	409	427	4.00
5742	94074	043000	EMD	MET_4	640	512	5.00
3778	94074	044936	EMD	NOAA_11	409	779	4.00
5747	94074	053000	EMD	MET_4	640	512	5.00
4106	94074	054829	EMD	NOAA_12	409	832	4.00
6782	94074	055400	EMD	DMSP_F11	1465	905	2.70
5752	94074	063000	EMD	MET_4	640	512	5.00
6777	94074	072600	EMD	DMSP_F10	1465	992	2.70
4098	94074	072839	EMD	NOAA_12	409	780	4.00
5757	94074	073000	EMD	MET_4	640	512	5.00
5762	94074	083000	EMD	MET_4	640	512	5.00
5767	94074	093000	EMD	MET_4	640	512	5.00
5772	94074	103000	EMD	MET_4	640	512	5.00
5777	94074	113000	EMD	MET_4	640	512	5.00
5782	94074	123000	EMD	MET_4	640	512	5.00
3786	94074	124856	EMD	NOAA_11	409	636	4.00
5787	94074	133000	EMD	MET_4	640	512	5.00
3794	94074	142745	EMD	NOAA_11	409	1029	4.00
5792	94074	143000	EMD	MET_4	640	512	5.00
4114	94074	152555	EMD	NOAA_12	409	378	4.00
5797	94074	153000	EMD	MET_4	640	512	5.00
6787	94074	153200	EMD	DMSP_F11	1465	1150	2.70
3802	94074	161051	EMD	NOAA_11	409	788	4.00
5802	94074	163000	EMD	MET_4	640	512	5.00
4122	94074	170253	EMD	NOAA_12	409	1031	4.00
5807	94074	173000	EMD	MET_4	640	512	5.00
5812	94074	183000	EMD	MET_4	640	512	5.00

ENTRY	DATE	TIME	ROI	SATELLITE	ELES	LINES	RESLN
4130	94074	184500	EMD	NOAA_12	409	806	4.00
5816	94074	193000	EMD	MET_4	640	512	5.00
6792	94074	193700	EMD	DMSP_F10	1465	1066	2.70
6797	94074	201700	EMD	DMSP_F10	1465	982	2.70
5821	94074	203000	EMD	MET_4	640	512	5.00
5825	94074	213000	EMD	MET_4	640	512	5.00
5829	94074	223000	EMD	MET_4	640	512	5.00
5833	94074	233000	EMD	MET_4	640	512	5.00
3810	94075	011627	EMD	NOAA_11	409	500	4.00
5837	94075	013000	EMD	MET_4	640	512	5.00
5841	94075	023000	EMD	MET_4	640	512	5.00
3818	94075	025636	EMD	NOAA_11	409	832	4.00
6812	94075	030000	EMD	DMSP_F11	1465	931	2.70
5845	94075	033000	EMD	MET_4	640	512	5.00
5849	94075	043000	EMD	MET_4	640	512	5.00
3826	94075	043720	EMD	NOAA_11	409	799	4.00
4138	94075	052649	EMD	NOAA_12	409	846	4.00
5854	94075	053000	EMD	MET_4	640	512	5.00
6817	94075	054100	EMD	DMSP_F11	1465	932	2.70
5859	94075	063000	EMD	MET_4	640	512	5.00
4146	94075	070707	EMD	NOAA_12	409	782	4.00
5864	94075	073000	EMD	MET_4	640	512	5.00
5869	94075	083000	EMD	MET_4	640	512	5.00
5874	94075	093000	EMD	MET_4	640	512	5.00
6802	94075	093500	EMD	DMSP_F10	1465	901	2.70
5879	94075	103000	EMD	MET_4	640	512	5.00
5884	94075	113000	EMD	MET_4	640	512	5.00
5889	94075	123000	EMD	MET_4	640	512	5.00
3834	94075	123726	EMD	NOAA_11	409	482	4.00
5894	94075	133000	EMD	MET_4	640	512	5.00
3842	94075	141525	EMD	NOAA_11	409	1038	4.00
6822	94075	141900	EMD	DMSP_F11	1465	1107	2.70
5899	94075	143000	EMD	MET_4	640	512	5.00
5904	94075	153000	EMD	MET_4	640	512	5.00
3850	94075	155818	EMD	NOAA_11	409	805	4.00
6827	94075	160000	EMD	DMSP_F11	1465	977	2.70
5909	94075	163000	EMD	MET_4	640	512	5.00
4154	94075	164116	EMD	NOAA_12	409	1046	4.00
5914	94075	173000	EMD	MET_4	640	512	5.00
4162	94075	182247	EMD	NOAA_12	409	864	4.00
5919	94075	183000	EMD	MET_4	640	512	5.00
5923	94075	193000	EMD	MET_4	640	512	5.00
5928	94075	203000	EMD	MET_4	640	512	5.00
6807	94075	204600	EMD	DMSP_F10	1465	962	2.70
5932	94075	213000	EMD	MET_4	640	512	5.00
5936	94075	223000	EMD	MET_4	640	512	5.00
5940	94075	233000	EMD	MET_4	640	512	5.00

ENTRY	DATE	TIME	ROI	SATELLITE	ELES	LINES	RESLN
3866	94076	010427	EMD	NOAA_11	409	224	4.00
5944	94076	013000	EMD	MET_4	640	512	5.00
5948	94076	023000	EMD	MET_4	640	512	5.00
3858	94076	024415	EMD	NOAA_11	409	848	4.00
5952	94076	033000	EMD	MET_4	640	512	5.00
6832	94076	034700	EMD	DMSP_F11	1465	934	2.70
3874	94076	042505	EMD	NOAA_11	409	801	4.00
6837	94076	042800	EMD	DMSP_F11	1465	885	2.70
5956	94076	043000	EMD	MET_4	640	512	5.00
4170	94076	050515	EMD	NOAA_12	409	837	4.00
5961	94076	053000	EMD	MET_4	640	512	5.00
5966	94076	063000	EMD	MET_4	640	512	5.00
4178	94076	064540	EMD	NOAA_12	409	765	4.00
5971	94076	073000	EMD	MET_4	640	512	5.00
6842	94076	080300	EMD	DMSP_F10	1465	925	2.70
5976	94076	083000	EMD	MET_4	640	512	5.00
5981	94076	093000	EMD	MET_4	640	512	5.00
5986	94076	103000	EMD	MET_4	640	512	5.00
5991	94076	113000	EMD	MET_4	640	512	5.00
3882	94076	122552	EMD	NOAA_11	409	320	4.00
5996	94076	123000	EMD	MET_4	640	512	5.00
6001	94076	133000	EMD	MET_4	640	512	5.00
3890	94076	140308	EMD	NOAA_11	409	1053	4.00
6006	94076	143000	EMD	MET_4	640	512	5.00
6011	94076	153000	EMD	MET_4	640	512	5.00
3898	94076	154543	EMD	NOAA_11	409	827	4.00
4186	94076	161947	EMD	NOAA_12	409	984	4.00
6016	94076	163000	EMD	MET_4	640	512	5.00
6872	94076	164700	EMD	DMSP_F11	1465	999	2.70
6021	94076	173000	EMD	MET_4	640	512	5.00
4194	94076	180024	EMD	NOAA_12	409	950	4.00
6026	94076	183000	EMD	MET_4	640	512	5.00
6847	94076	191300	EMD	DMSP_F10	1465	1073	2.70
6030	94076	193000	EMD	MET_4	640	512	5.00
6035	94076	203000	EMD	MET_4	640	512	5.00
6039	94076	213000	EMD	MET_4	640	512	5.00
6852	94076	215500	EMD	DMSP_F10	1465	935	2.70
6043	94076	223000	EMD	MET_4	640	512	5.00
6047	94076	233000	EMD	MET_4	640	512	5.00
6051	94077	013000	EMD	MET_4	640	512	5.00
6055	94077	023000	EMD	MET_4	640	512	5.00
3906	94077	023157	EMD	NOAA_11	409	853	4.00
6059	94077	033000	EMD	MET_4	640	512	5.00
3914	94077	041253	EMD	NOAA_11	409	793	4.00
6063	94077	043000	EMD	MET_4	640	512	5.00
4202	94077	044348	EMD	NOAA_12	409	811	4.00
6068	94077	053000	EMD	MET_4	640	512	5.00
4210	94077	062405	EMD	NOAA_12	409	757	4.00
6073	94077	063000	EMD	MET_4	640	512	5.00

ENTRY	DATE	TIME	ROI	SATELLITE	ELES	LINES	RESLN
6078	94077	073000	EMD	MET_4	640	512	5.00
6083	94077	083000	EMD	MET_4	640	512	5.00
6857	94077	083200	EMD	DMSP_F10	1465	986	2.70
6862	94077	091100	EMD	DMSP_F10	1465	882	2.70
6088	94077	093000	EMD	MET_4	640	512	5.00
6093	94077	103000	EMD	MET_4	640	512	5.00
6098	94077	113000	EMD	MET_4	640	512	5.00
6103	94077	123000	EMD	MET_4	640	512	5.00
6108	94077	133000	EMD	MET_4	640	512	5.00
3922	94077	135052	EMD	NOAA_11	409	1055	4.00
6113	94077	143000	EMD	MET_4	640	512	5.00
6118	94077	153000	EMD	MET_4	640	512	5.00
3938	94077	153304	EMD	NOAA_11	409	864	4.00
4226	94077	155852	EMD	NOAA_12	409	832	4.00
6123	94077	163000	EMD	MET_4	640	512	5.00
6877	94077	163300	EMD	DMSP_F11	1465	1025	2.70
6882	94077	171500	EMD	DMSP_F11	1465	956	2.70
6128	94077	173000	EMD	MET_4	640	512	5.00
4234	94077	173823	EMD	NOAA_12	409	998	4.00
6133	94077	183000	EMD	MET_4	640	512	5.00
6137	94077	193000	EMD	MET_4	640	512	5.00
6887	94077	194300	EMD	DMSP_F10	1465	1058	2.70
6892	94077	202200	EMD	DMSP_F10	1465	971	2.70
6142	94077	203000	EMD	MET_4	640	512	5.00
6146	94077	213000	EMD	MET_4	640	512	5.00
6150	94077	223000	EMD	MET_4	640	512	5.00
6154	94077	233000	EMD	MET_4	640	512	5.00
6158	94078	013000	EMD	MET_4	640	512	5.00
3946	94078	021940	EMD	NOAA_11	409	854	4.00
6162	94078	023000	EMD	MET_4	640	512	5.00
6166	94078	033000	EMD	MET_4	640	512	5.00
3954	94078	040042	EMD	NOAA_11	409	783	4.00
4242	94078	042229	EMD	NOAA_12	409	767	4.00
6170	94078	043000	EMD	MET_4	640	512	5.00
6175	94078	053000	EMD	MET_4	640	512	5.00
4258	94078	060218	EMD	NOAA_12	409	814	4.00
6180	94078	063000	EMD	MET_4	640	512	5.00
6897	94078	070000	EMD	DMSP_F10	1465	905	2.70
6185	94078	073000	EMD	MET_4	640	512	5.00
4250	94078	074222	EMD	NOAA_12	409	646	4.00
6190	94078	083000	EMD	MET_4	640	512	5.00
6195	94078	093000	EMD	MET_4	640	512	5.00
6902	94078	094000	EMD	DMSP_F10	1465	896	2.70
6200	94078	103000	EMD	MET_4	640	512	5.00
6205	94078	113000	EMD	MET_4	640	512	5.00
6210	94078	123000	EMD	MET_4	640	512	5.00
6215	94078	133000	EMD	MET_4	640	512	5.00
3962	94078	133836	EMD	NOAA_11	409	1044	4.00

ENTRY	DATE	TIME	ROI	SATELLITE	ELES	LINES	RESLN
6220	94078	143000	EMD	MET_4	640	512	5.00
6917	94078	152000	EMD	DMSP_F11	1465	1014	2.70
3970	94078	152021	EMD	NOAA_11	409	909	4.00
6225	94078	153000	EMD	MET_4	640	512	5.00
4282	94078	153848	EMD	NOAA_12	409	554	4.00
6230	94078	163000	EMD	MET_4	640	512	5.00
6922	94078	170200	EMD	DMSP_F11	1465	967	2.70
4266	94078	171638	EMD	NOAA_12	409	1023	4.00
6235	94078	173000	EMD	MET_4	640	512	5.00
6907	94078	181100	EMD	DMSP_F10	1465	932	2.70
6240	94078	183000	EMD	MET_4	640	512	5.00
4274	94078	185901	EMD	NOAA_12	409	785	4.00
6244	94078	193000	EMD	MET_4	640	512	5.00
6249	94078	203000	EMD	MET_4	640	512	5.00
6912	94078	205100	EMD	DMSP_F10	1465	956	2.70
6253	94078	213000	EMD	MET_4	640	512	5.00
6257	94078	223000	EMD	MET_4	640	512	5.00
6261	94078	233000	EMD	MET_4	640	512	5.00
6265	94079	013000	EMD	MET_4	640	512	5.00
3978	94079	020726	EMD	NOAA_11	409	846	4.00
6269	94079	023000	EMD	MET_4	640	512	5.00
6273	94079	033000	EMD	MET_4	640	512	5.00
3986	94079	034831	EMD	NOAA_11	409	771	4.00
4290	94079	040119	EMD	NOAA_12	409	261	4.00
6277	94079	043000	EMD	MET_4	640	512	5.00
6282	94079	053000	EMD	MET_4	640	512	5.00
4298	94079	054033	EMD	NOAA_12	409	837	4.00
6927	94079	062900	EMD	DMSP_F10	1465	857	2.70
6287	94079	063000	EMD	MET_4	640	512	5.00
4306	94079	072046	EMD	NOAA_12	409	784	4.00
6292	94079	073000	EMD	MET_4	640	512	5.00
6932	94079	080800	EMD	DMSP_F10	1465	919	2.70
6297	94079	083000	EMD	MET_4	640	512	5.00
6302	94079	093000	EMD	MET_4	640	512	5.00
6307	94079	103000	EMD	MET_4	640	512	5.00
6937	94079	104800	EMD	DMSP_F10	1465	918	2.70
6312	94079	113000	EMD	MET_4	640	512	5.00
6317	94079	123000	EMD	MET_4	640	512	5.00
3994	94079	132625	EMD	NOAA_11	409	1002	4.00
6322	94079	133000	EMD	MET_4	640	512	5.00
6327	94079	143000	EMD	MET_4	640	512	5.00
4002	94079	150735	EMD	NOAA_11	409	964	4.00

ENTRY	DATE	TIME	ROI	SATELLITE	ELES	LINES	RESLN
6962	94079	150800	EMD	DMSP_F11	1465	980	2.70
4322	94079	151828	EMD	NOAA_12	409	273	4.00
6332	94079	153000	EMD	MET_4	640	512	5.00
6337	94079	163000	EMD	MET_4	640	512	5.00
4314	94079	165458	EMD	NOAA_12	409	1041	4.00
6342	94079	173000	EMD	MET_4	640	512	5.00
6347	94079	183000	EMD	MET_4	640	512	5.00
4330	94079	183655	EMD	NOAA_12	409	780	4.00
6942	94079	191900	EMD	DMSP_F10	1465	1010	2.70
6351	94079	193000	EMD	MET_4	640	512	5.00
6356	94079	203000	EMD	MET_4	640	512	5.00
6360	94079	213000	EMD	MET_4	640	512	5.00
6364	94079	223000	EMD	MET_4	640	512	5.00
6368	94079	233000	EMD	MET_4	640	512	5.00
6372	94080	013000	EMD	MET_4	640	512	5.00
6695	94080	015515	EMD	NOAA_11	409	828	4.00
6376	94080	023000	EMD	MET_4	640	512	5.00
6380	94080	033000	EMD	MET_4	640	512	5.00
6687	94080	033616	EMD	NOAA_11	409	765	4.00
6384	94080	043000	EMD	MET_4	640	512	5.00
4338	94080	051855	EMD	NOAA_12	409	844	4.00
6389	94080	053000	EMD	MET_4	640	512	5.00
6394	94080	063000	EMD	MET_4	640	512	5.00
4346	94080	065916	EMD	NOAA_12	409	778	4.00
6399	94080	073000	EMD	MET_4	640	512	5.00
6404	94080	083000	EMD	MET_4	640	512	5.00
6952	94080	083700	EMD	DMSP_F10	1465	980	2.70
6957	94080	091700	EMD	DMSP_F10	1465	889	2.70
6409	94080	093000	EMD	MET_4	640	512	5.00
6414	94080	103000	EMD	MET_4	640	512	5.00
6419	94080	113000	EMD	MET_4	640	512	5.00
6424	94080	123000	EMD	MET_4	640	512	5.00
6703	94080	131405	EMD	NOAA_11	409	973	4.00
6429	94080	133000	EMD	MET_4	640	512	5.00
6434	94080	143000	EMD	MET_4	640	512	5.00
6711	94080	145458	EMD	NOAA_11	409	1003	4.00
6439	94080	153000	EMD	MET_4	640	512	5.00
6972	94080	155500	EMD	DMSP_F11	1465	1084	2.70
6444	94080	163000	EMD	MET_4	640	512	5.00
4354	94080	163325	EMD	NOAA_12	409	1024	4.00
6449	94080	173000	EMD	MET_4	640	512	5.00
6977	94080	173500	EMD	DMSP_F11	1465	998	2.70
4362	94080	181439	EMD	NOAA_12	409	891	4.00

ENTRY	DATE	TIME	ROI	SATELLITE	ELES	LINES	RESLN
6454	94080	183000	EMD	MET_4	640	512	5.00
6458	94080	193000	EMD	MET_4	640	512	5.00
6982	94080	194800	EMD	DMSP_F10	1465	1042	2.70
6987	94080	202800	EMD	DMSP_F10	1465	958	2.70
6463	94080	203000	EMD	MET_4	640	512	5.00
6467	94080	213000	EMD	MET_4	640	512	5.00
6471	94080	223000	EMD	MET_4	640	512	5.00
6475	94080	233000	EMD	MET_4	640	512	5.00
6479	94081	013000	EMD	MET_4	640	512	5.00
6719	94081	014306	EMD	NOAA_11	409	806	4.00
6483	94081	023000	EMD	MET_4	640	512	5.00
7012	94081	032200	EMD	DMSP_F11	1465	929	2.70
6735	94081	032356	EMD	NOAA_11	409	784	4.00
6487	94081	033000	EMD	MET_4	640	512	5.00
6491	94081	043000	EMD	MET_4	640	512	5.00
4370	94081	045723	EMD	NOAA_12	409	830	4.00
7017	94081	050300	EMD	DMSP_F11	1465	905	2.70
6727	94081	050426	EMD	NOAA_11	409	464	4.00
6496	94081	053000	EMD	MET_4	640	512	5.00
6501	94081	063000	EMD	MET_4	640	512	5.00
6992	94081	070600	EMD	DMSP_F10	1465	906	2.70
6506	94081	073000	EMD	MET_4	640	512	5.00
6511	94081	083000	EMD	MET_4	640	512	5.00
6516	94081	093000	EMD	MET_4	640	512	5.00
6997	94081	094500	EMD	DMSP_F10	1465	891	2.70
6521	94081	103000	EMD	MET_4	640	512	5.00
6526	94081	113000	EMD	MET_4	640	512	5.00
6531	94081	123000	EMD	MET_4	640	512	5.00
6743	94081	130245	EMD	NOAA_11	409	815	4.00
6536	94081	133000	EMD	MET_4	640	512	5.00
6541	94081	143000	EMD	MET_4	640	512	5.00
6751	94081	144236	EMD	NOAA_11	409	1015	4.00
6546	94081	153000	EMD	MET_4	640	512	5.00
7022	94081	154100	EMD	DMSP_F11	1465	1121	2.70
4378	94081	161153	EMD	NOAA_12	409	961	4.00
7027	94081	162100	EMD	DMSP_F11	1465	1145	2.70
6551	94081	163000	EMD	MET_4	640	512	5.00
6556	94081	173000	EMD	MET_4	640	512	5.00
4386	94081	175211	EMD	NOAA_12	409	987	4.00
7002	94081	181600	EMD	DMSP_F10	1465	1002	2.70
6561	94081	183000	EMD	MET_4	640	512	5.00
6565	94081	193000	EMD	MET_4	640	512	5.00
6570	94081	203000	EMD	MET_4	640	512	5.00
7007	94081	205600	EMD	DMSP_F10	1465	1040	2.70
6574	94081	213000	EMD	MET_4	640	512	5.00
6578	94081	223000	EMD	MET_4	640	512	5.00
7032	94082	055000	EMD	DMSP_F11	1465	914	2.70
7200	94082	013000	EMD	MET_4	640	512	5.00
7204	94082	023000	EMD	MET_4	640	512	5.00
7208	94082	033000	EMD	MET_4	640	512	5.00

7212	94082	043000	EMD	MET_4	640	512	5.00
7217	94082	053000	EMD	MET_4	640	512	5.00
7222	94082	063000	EMD	MET_4	640	512	5.00
7227	94082	073000	EMD	MET_4	640	512	5.00

Appendix B

Archive Data Format Descriptions

By Data Processing Level

Level 1: Satellite Image Files

Satellite image filenames as they appear on tape have the following naming convention:

SSS_CCC_ROI_DDD_HH.Tif

where

SSS - Satellite identifier:

F10	DMSP F-10
F11	DMSP F-11
N11	NOAA-11
N12	NOAA-12
M04	METEOSAT-4

CCC - spectral channel identifier
ROI - Region of Interest:
EMD for Eastern Mediterranean and Desert Area
DDD - Julian day for which the image data are valid
HH - UTC hour of image data
Tif - TIF file format

File and Record Structure

All image files contain fixed-length records. The number of lines and number of elements in an image file are contained in the Related Entries (RE) SDB information file that is provided with the tape, under the heading of SATIMG:

NUM_LINES	Number of image data lines in the file.
ELEM_PER_LINE	Number of elements (pixels) per line.
BYTES_PER_ELEMENT	Number of bytes per pixel. This number is 1 for all SERCAA imager sensor data.

Image file data are stored in Tagged Image File Format (TIF), therefore an alternative way to determine image dimensions is to read the TIF header and examine the width and height fields.

Image pixel values represent either counts or albedo for visible data, and brightness temperatures for thermal infrared data. Table B-1 summarizes the attributes of the SERCAA image data values.

Table B-1 Satellite image characteristics

Satellite ID (SSS)	Spectral Channel (CCC)	Channel Type	Wavelength Band	Physical Value
F10 or F11	001	Visible	0.4 - 1.1 µm	Counts[1]
	002	Long-Wave IR	10 -12 µm	Brightness Temp.[2]
N11 or N12	001	Visible	0.63 µm	Albedo[3]
	002	Near-IR	0.86 µm	Albedo
	003	Mid-Wave IR	3.7 µm	Brightness Temp.
	004	Long-Wave IR	10.7 µm	Brightness Temp.
	005	Long-Wave IR	11.8 µm	Brightness Temp.
M04	001	Visible	0.5 - 0.75 µm	Counts
	002	Long-Wave IR	10.5 - 12.5 µm	Brightness Temp.

[1] Visible counts range from 0 - 255. High counts denote highly reflective surfaces and low counts denote poorly reflective surfaces.

[2] Brightness temperatures are byte-encoded such that the range 0 - 255 corresponds to the temperature range 327.5 K to 200.0 K. The relation between byte values and temperature is linear over this range; the conversion from byte value B to brightness temperature T is given by the relation:

$$T = -0.5\, B + 327.5.$$

[3] Albedo values are byte-encoded such that the range 0 - 255 corresponds to the albedo range 0 - 100%. The relation between byte values and percent albedo is linear; the conversion from byte value B to percent albedo A is given by the relation

$$A = 0.392\, B.$$

Level 1: Latitude-Longitude File

Latitude-longitude filenames as they appear on tape have the following naming convention:

SSS_LAT_ROI_DDD_HH.Dat

where

SSS - Satellite identifier:

F10	DMSP F-10
F11	DMSP F-11
N11	NOAA-11
N12	NOAA-12
M04	METEOSAT-4

LAT - a constant that identifies the file as a latitude-longitude file

ROI - Region of Interest for which the latitude-longitude file is valid:

EMD for Eastern Mediterranean and Desert Area

DDD - Julian day of satellite data for which the Earth locations are valid

HH - UTC hour of the satellite data for which the Earth locations are valid

File and Record Structure

Latitude-longitude Earth location files contain fixed-length records, the number and size of which depend on both the size of the corresponding image files and the satellite type. There is always one latitude-longitude record corresponding to each satellite image file record, where a satellite image file record contains one image scan line of information. However, latitude-longitude data are subsampled, relative to the sensor data, along a scan line. There is one latitude-longitude pair for every n^{th} image pixel, where n is a function of satellite. A linear interpolation is used to compute Earth location for intermediate pixels between latitude-longitude reference points.

The information necessary for interpreting a latitude-longitude file record is contained in the Related Entries (RE) SDB information file provided with the tape, under the heading of LATLON:

LL_REC_LEN	Record length in bytes.
LL_LINE_INTERVAL	The number of image file records per lat-lon record. For the March 1994 data set this number is always 1.
LL_ELEM_INTERVAL	The subsampling rate of lat-lon information relative to the corresponding satellite data. For example, if LL_ELEM_INTERVAL = 40, there is one latitude-longitude pair for every 40th image pixel in the scan line (i.e., for pixels 1, 41, 81, ...). Linear interpolation is required to retrieve Earth location information for intermediate pixels 2-40, 42-80, ...
LL_ELEM_PER_LINE	This is the number of latitude-longitude elements per latitude-longitude file record.

A latitude-longitude file data element is a 4-byte structure that contains the scaled latitude and longitude for a given pixel. Thus the length of a latitude-longitude file record in bytes is given by:

LL_REC_LEN = 4 * LL_ELEM_PER_LINE

where the 4 bytes consist of two 16-bit integer variables: LONG and LAT. The storage convention is as follows:

LONG	Pixel longitude * 128. To obtain the floating-point longitude, FLONG = LONG / 128. Longitude range is -180° to 180°, positive east.
LAT	Pixel latitude * 128. to obtain floating-point latitude, FLAT = LAT / 128. Latitude range is -90° to 90°, positive north.

Level 1: **Angles File**

The angles filenames as they appear on tape have the following naming convention:

SSS_ANG_ROI_DDD_HH.Dat

where

SSS - Satellite identifier:

F10	DMSP F-10
F11	DMSP F-11
N11	NOAA-11
N12	NOAA-12
M04	METEOSAT-4

ANG - a constant that identifies the file as an angles file

ROI - Region of Interest for which the angles file is valid:

EMD for Eastern Mediterranean and Desert Area

DDD - Julian day of satellite data for which the angles are valid

HH - UTC hour of the satellite data for which the angles are valid

File and Record Structure

Angle files contain fixed-length records, the number and size of which depend on both the size of the corresponding image files and the satellite type. There is always one angles record corresponding to each satellite image file record, where a satellite image file record contains one image scan line of information. However, angle data are subsampled, relative to the sensor data, along a scan line. There is one set of angles for every n^{th} image pixel, where n is a function of satellite. A linear interpolation is used to compute angle values for intermediate pixels between angle reference points.

The information necessary for interpreting an angles file record is contained in the Related Entries (RE) SDB information file provided with the tape, under the heading of ANGLES:

ANG_REC_LEN	Record length in bytes.
ANG_LINE_INTERVAL	The number of image file records per angles record. This number is almost always 1.
ANG_ELEM_INTERVAL	The subsampling rate of angles information relative to the corresponding satellite image. For example, if ANG_ELEM_INTERVAL = 8, there is one set of angles valid for every eighth image pixel in the scan line (i.e., for pixels 1, 9, 17, 25, ...). Linear interpolation is required to retrieve angles information for intermediate pixels 2-8, 10-16, 18-24, ...
ANG_ELEM_PER_LINE	This is the number of angles elements per angles file record.

An angles file data element is a 12-byte structure containing three angles that define the satellite and solar viewing geometry for a given pixel. Thus the length of an angles file record in bytes is given by:

ANG_REC_LEN = 12 * ANG_ELEM_PER_LINE

where the 12 bytes consist of three 32-bit floating-point variables: SATZEN, SOLZEN, and AZIMUTH corresponding to the satellite zenith, the solar zenith, and the satellite/solar azimuth angles respectively (Figure B-1). Note: Angle files were generated on a VMS computer. To interpret these floating-point numbers on a UNIX machine it is necessary to convert from VMS to IEEE floating-point formats. Most UNIX operating systems provide a utility to perform this conversion. Angle measurement conventions are as follows:

SATZEN	Scene satellite zenith angle, 0^{o} - 90^{o}.
SOLZEN	Scene solar zenith angle, 0^{o} - 180^{o}.
AZIMUTH	Relative angle between the solar and satellite azimuth angles, 0^{o} - 359^{o}. When AZIMUTH = 0^{o}, the sun is directly behind the satellite (i.e., the viewed point, the satellite, and the sun are collinear). When AZIMUTH = 180^{o}, the satellite is looking directly into the sun (the satellite squints to compensate).

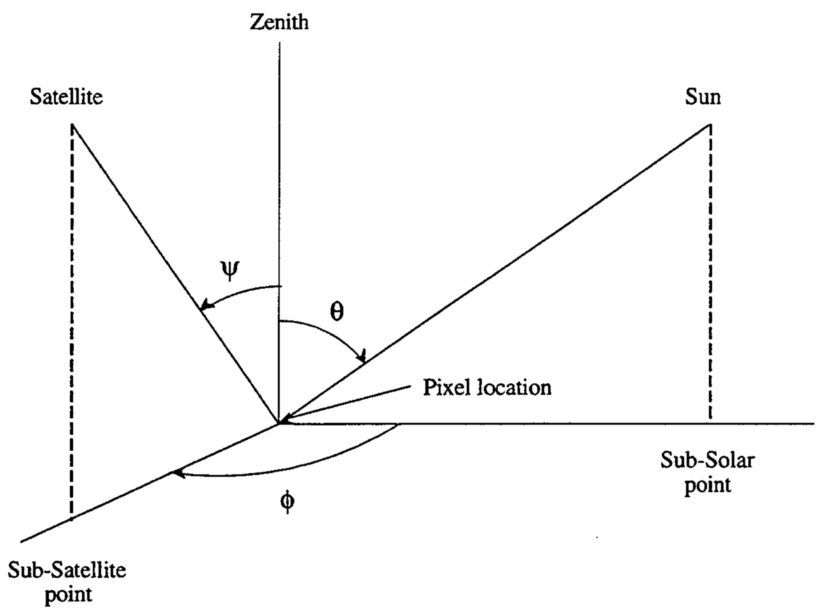


Figure B-1. Satellite-Earth-Solar Geometry (after Taylor and Stowe, 1984)

ψ - satellite zenith angle

θ - solar zenith angle

ϕ - sun-satellite azimuth angle

Level 2: **Nephanalysis Products**

Nephanalysis products are stored as bit-encoded byte values known as MCF (cloud Mask and Confidence Flag). MCF filenames as they appear on tape have the following naming convention:

SSS_MCF_ROI_DDD_HH.Dat or .Tif

where

SSS - Satellite identifier:
- F10 DMSP F-10
- F11 DMSP F-11
- N11 NOAA-11
- N12 NOAA-12
- M04 METEOSAT-4

MCF - a constant that identifies the file as an MCF file
ROI - Region of Interest for which the product is valid:
EMD for Eastern Mediterranean and Desert Area
DDD - Julian day for which the product is valid
HH - UTC hour for which the product is valid
Dat - Raw product file format
Tif - TIF file format

File and Record Structure

Level 2 processing is performed on square arrays of image pixels, therefore the size of the resultant MCF product files is an integral number of the analysis array size. MCF files contain fixed-length records, the number and size of which depends on both the size of the corresponding image files and the satellite type. The following table specifies how to determine the record size and number of records in an MCF file. Let NCOLS and NROWS be the number of columns and rows, respectively, in the corresponding satellite image file; then:

If the image satellite id is:	Then the MCF file record size is:	And the number of lines is:
DMSP F10 or F11	NCOLS - MOD(NCOLS, 16)	NROWS - MOD(NROWS, 16)
NOAA 11 or 12	NCOLS - MOD(NCOLS, 16)	NROWS - MOD(NROWS, 16)
METEOSAT 4	See Associated RE File or TIF Header	See Associated RE File or TIF Header

where MOD is the FORTRAN modulus function (e.g., if an F10 pass has 1465 columns per scan line, then the MCF record size is 1456). The MCF file is stored in Tagged Image File Format (TIF), therefore an alternative way to determine file dimensions is to read the TIF header and examine the width and height fields.

The format of an MCF file is the same regardless of the satellite platform it was derived from. The first byte of the first record of the MCF file corresponds to the first byte of the first record in the corresponding image data file. Across each scan line there is a one-to-one correspondence between the image and MCF files out to the number of bytes computed above for each record. As can be seen in the above table, the MCF and image file sizes are not always the same. However, the two files are always aligned with respect to the upper-left corner of each.

There is one 8-bit MCF byte per analyzed image pixel. MCF bytes are bit-packed according to the following convention:

Bit 0 (least significant) is the cloud/no-cloud bit. If bit 0 is off, the corresponding image pixel is clear; if bit 0 is on, it is completely cloudy.

Bit 1 is the low cloud bit. If bit 1 is on, the pixel contains low cloud as determined by an appropriate spectral (or other) signature test.

Bit 2 is the thin cirrus cloud bit. If bit 2 is on, the pixel contains cirrus as determined by an appropriate spectral (or other) signature test.

Bit 3 is the precipitating or cumulonimbus cloud bit. If bit 3 is on, the pixel contains clouds with large vertical extent.

Bit 4 is the partly cloudy bit. If bit 4 is on, the pixel is partly cloudy. If bit 4 is on, bit 0 is off. Not used for this data set.

Bit 5 is the bad data bit. It is set whenever satellite data are missing or unreliable. If set, all other bits should be ignored.

Bits 6 and 7 contain the confidence level attached to the accuracy of the cloud/no-cloud decision for the corresponding cloudy image pixel. Confidence levels are rated as 0 for missing data, 1 for low confidence, 2 for mid-level confidence, and 3 for high confidence.

Low cloud, thin cirrus, and cumulonimbus conditions are always associated with completely cloudy conditions (i.e., bit 0 will always be on in the presence of one or more of these conditions). Cloud level and cloud type are not detected under partly cloudy conditions (i.e., if bit 4 is on, bits 1 through 3 will be off).

Example:

MCF byte	1	1	0	0	0	1	0	1	(C5 in hex)
bit position	7	6	5	4	3	2	1	0	

The corresponding image pixel is classified as cloud covered (bit 0) with thin cirrus (bit 2) that has been detected with a high level of confidence (bits 6 and 7).

Level 3: Layered Product

The layered product filename as it appears on tape has the following naming convention:

SAT_LYR_ROI_DDD_HH.DAT

where:

SAT - Satellite identifier:
- F10 DMSP F-10
- F11 DMSP F-11
- N11 NOAA-11
- N12 NOAA-12
- M04 METEOSAT-4

LYR is a constant that denotes the file is a layered product
ROI - Region of Interest:
EMD for Eastern Mediterranean and Desert Area
DDD - Julian day
HH - GMT hour

File Structure

The layered product file contains 20349 (153 rows x 133 columns) record structures.

Record Structure

Each record contains data values valid for one grid point within a 153 (rows) by 133 (columns) two-dimensional grid. The grid is superimposed on a hemispheric secant polar stereographic map projection. Grid resolution is based on a whole mesh grid spacing of 381 km at 60 degrees latitude and nested grids are defined in terms of the number of grid cells that fit within a whole mesh grid. The layered product grid is a 1/16th mesh grid (i.e., 16 by 16 grid cells per whole mesh box.)

Table B-2 summarizes the contents of each record. Figure B-2 contains the C data structure that was used to create the data file.

Important information for reading the layering output file

The layering output file was written using an "fwrite" statement with a byte length field of "sizeof(LAYER_OUTPUT)" where LAYER_OUTPUT is the output structure contained in Figure B-2. Note that while the sum of the structure elements in Table B-2 is 59 bytes the sizeof(LAYER_OUTPUT) operator evaluates to 64 bytes. The reason for the 5 byte discrepancy is the way word alignment if performed under the UNIX operating system. The beginning of fields 6-9 and 11-14 in Table B-2 are automatically word aligned causing an additional 4 bytes to be added to the structure. The remaining byte to account for the 5 byte difference comes from rounding up the odd number of bytes in the structure. Thus, to read this file it is necessary to either use the sizeof operator in the read statement (preferred) or to hard-wire a value of 64 bytes.

Table B-2: Layered Product Record Structure

Field	Description	Units	Range	Missing or bad value	Byte length
1	Absolute 16th-mesh row number (i)		1-1024		2
2	Absolute 16th-mesh column number (j)		1-1024		2
3	SDB IR entry number			0	2
4	Julian day (ddd)		1-366		4
5	UTC (hhmm)		0-2359		2
6-9	Cloud temperature variance for each layer	GS*100			8
10	# pixels in grid box			0	2
11-14	# pixels in each layer			0	8
15-18	Cloud top temperature for each layer	GS*100			4
19-22	Cloud type for each layer*		1-2		4
23-26	# low cloud pixels in each layer				4
27-30	# thin cirrus pixels in each layer				4
31-34	# precipitating-cloud pixels in each layer				4
35-38	Mean confidence level for each layer		10-30		4
39	Sunrise time		0-235		1
40	Sunset time		0-235		1
41	Satellite platform ID				1
42	# data dropouts in grid box				1
43	# partially cloud-filled pixels				1

*Cloud type not computed

```
/* Layering output structure

   Daniel Peduzzi (AER) 9/27/94
     structure content by Robert P. d'Entremont (AER) 9/1994
*/

#ifndef NCLASSES
#  define NCLASSES (4)
#endif

#ifndef _LAYER_OUTPUT
#define _LAYER_OUTPUT

#define BYTE unsigned char

typedef struct {

  short i;                              /* 16th-mesh absolute row (1-1024)               */
  short j;                              /* 16th-mesh absolute column (1-1024)            */

  short sdb_ir_entry;                   /* SDB entry number corresponding to IR data     */
  int yyddd;                            /* Sensor data Julian day                        */
  short hhmm;                           /* Sensor data valid time (UTC) hhmm             */

  short layer_var[NCLASSES];            /* Temperature variance*100 for cloud layer i    */
  short num_pixels;                     /* Total # of pixels in 16th-mesh box            */
  short n_layer_pix[NCLASSES];          /* Total # pixels in layer i                     */
  BYTE meantemp[NCLASSES];              /* Mean cloud top temperature for layer i        */
  BYTE cloud_type[NCLASSES];            /* Cloud type for layer i (1 or 2)               */
  BYTE low_cloud[NCLASSES];             /* # low cloud pixels in layer i                 */
  BYTE thin_cirrus[NCLASSES];           /* # thin cirrus pixels in layer i               */
  BYTE precip[NCLASSES];                /* # precipitating-cloud pixels in layer i       */
  BYTE confidence[NCLASSES];            /* Confidence for layer i (10-30)                */

  BYTE sunrise;                         /* Sunrise time (UTC) (0-235)                    */
  BYTE sunset;                          /* Sunset time (UTC) (0-235)                     */
  BYTE vid;                             /* Satellite vehicle (platform) ID               */
  BYTE dropouts;                        /* Total # of data dropouts in 16th-mesh box     */
  BYTE partial;                         /* Total # of partially-cloud-filled pixels      */

} LAYER_OUTPUT;

#undef BYTE

#endif
```

Figure B-2: Level 3 data structure

Level 4: **Integrated Product**

The integrated product filename as it appears on tape has the following naming convention:

ALL_IAN_ROI_DDD_HH.Dat

where

ALL and IAN are constants (Integrated ANalysis from ALL sensors)
ROI - Region of Interest for which the product is valid
Possible values:
EMD for Eastern Mediterranean and Desert Area
DDD - Julian day for which the integrated product is valid
HH - GMT hour for which the integrated product is valid

File Structure

The integrated product file contains 20,349 records (133 columns by 153 rows), each 64 bytes in length.

Record Structure

Each record contains data values valid for one grid point within a 153 (rows) X 133 (columns) 2-D grid. The grid is superimposed on a hemispheric secant polar stereographic map projection. Grid resolution is based on a whole-mesh grid spacing of 381 km at 60^{o} latitude and nested grids are defined in terms of the number of grid cells that fit within a whole mesh grid. The integrated product grid is a $1/16^{th}$ mesh grid (i.e., 16 X 16 cells per whole mesh box).

Table B-3 summarizes the contents of each record. All values are 16-bit integers and one grid cell occupies 64 bytes. Figure B-3 contains the C data structure used to create the output file.

Table B-3. Integrated Product Record Structure*

Field	Description	Units	Range	Missing or bad value	Comments
1	Absolute 16th-mesh column number (i)		227 - 451		
2	Absolute 16th-mesh row number (j)		13 - 395		
3	Number of cloud layers in (i,j)		0 - 4	-999	
4	Total cloud fraction for (i,j)	Percent	0 - 100	-999	
5-8	Cloud fraction by layer for (i,j)	Percent	0 - 100	-999	
9-12	Cloud top temperature by layer	K*10	2000-3275	-999	
13-16	Cloud top height by layer	Meters	0-13500	-999	
17-20	Cloud type by layer		0 - 9	-999	Not calculated.
21	Total cloud fraction error for (i,j)	Percent	0 - 100	-999	
22-25	Layer cloud fraction error for (i,j)	Percent	0 - 100	-999	
26-29	Layer confidence flags for (i,j)	Flag*10	10 - 30	-999	Discrete values for low to high confidence
30-32	Database entry numbers for input satellite analyses				Corresponds to directory names on tar tape

*all values are 16-bit integers

```
/* EMDA definitions */

#define NLINE 153
#define NCOL 133
#define NLYRS 4


#define MIN_I 731
#define MIN_J 353

typedef unsigned char byte;

/* integration output structure */

typedef struct {
   short i;                    /* absolute 16th mesh coord          */
   short j;
   short nlayers;              /* number of layers                  */
   short fraction;             /* total cloud fraction              */

   short lyr_frc[NLYRS];       /* layer cloud fraction              */
   short t_cld[NLYRS];         /* layer cloud top temp (K*10)       */
   short z_cld[NLYRS];         /* layer cloud top height (m)        */
   short cld_typ[NLYRS];       /* layer cloud type                  */

   short error;                /* total cloud amount error          */
   short lyr_err[NLYRS];       /* layer cloud amount error          */
   short conf[NLYRS];          /* layer confidence measure          */
   short sdb_entry[3];         /* input entry number(s)             */
} INTEGRATION;
```

Figure B-3: Integration output data structure

Appendix C

Data Extraction Guide

****SERCAA DATA SET RELEASE TO DNA****

****
What should I have ?

DNA_RELEASE.TXT	This document.
(2) 8 mm D8-112 tapes	One tape, labeled DNA MAR94 EMD IA/RE 071-078, contains the SERCAA Integrated Analysis (SIA) data files and the Related Entry(RE) data files for the first 8 days of processing. The other tape, labeled DNA MAR94 EMD IA/RE 079-080, contains the SIA and the RE data files for the last two days of processing. NOTE: a RE consists of Satellite, Latitude/Longitude, Angles(Geometry) and Product(cloud mask) data files.

****
What type of tape drive was used ?

A SUN Exabyte EXB-8500 8 mm tape drive recording in high density mode (5 gig).

****
What utility was used to create the release tapes ?

The data were placed on the tapes using a SUN SPARC II running SUN OS 4.1.2. The following tar command syntax was used:

```
sun% tar cvBf /dev/nrst8 somedirectory
```

****
How are the data arranged on the release tape ?

The data are placed on the tape as a series of tar files. The SIA data files are placed on the tape first and the RE data files are placed on the tape last.

The SIA tar files are placed on the tape such that each tar file represents a directory that contains all the SIA data for a particular day (day 94071 through day 94080). Each directory name follows the convention:

CYYJJJ

where:

C = century (9 for 19XX)
YY = year
JJJ = Julian day

A SIA file and SIA SDB information file exists for each hour that an analysis was performed. Each SIA file has been named using the following convention:

Positions 1-4	Platform:	
		all_ = All satellite platforms are used to create a SIA.
Positions 5-8	Type of file:	
		ian_ = integrated analysis file
		sdb_ = SERCAA data base (SDB) information file
Positions 9-12	Region of interest:	
		(Given in 16th-mesh coordinates)
		eas_ = East Asia Area (EASA). (i,j) = (227,13) to (451,395)
		can_ = Canada Area (CANA). (409,597) to (557,711)
		cns_ = Central, Northern South America Area (CNSA). (413,877) to (651,1011)
		emd_ = Eastern Mediterranean, Desert Area (EMDA). (731,353) to (863,505)
Positions 13-16	Julian day:	
		081_ = Julian day 081 etc. ...
Positions 17-18	Hour:	
		00 = SIA for hour 00 etc. ...
Positions 19-22	Extension:	
		.dat = raw-format file extension
Example:		all_ian_eas_081_10.dat

***** PLEASE NOTE THE CHANGE DOCUMENTED BELOW *****

The RE tars have changed from previous data releases. Previously each tar file represented a directory that contained all the related data used as input to create at least one of the SIA data files. This method produced a few hundred tar files on the tape, which resulted in excessive tape operations. To reduce the number of tape operations, there are now one or more directories per tar file.

To facilitate this change, a grouping size and a grouping number were utilized to distribute the data in an efficient and logical manner. The grouping size determines the maximum number of directories that will be grouped together in one tar file. The directories grouped together are identified by a grouping number. The grouping is implemented by using the grouping number as the parent directory of the group. The grouping number range is from 1 to (ND/GS) + 1 ,where ND is the number of related entry directories and GS is the grouping size. For example, this data release utilizes a grouping size of 10. Therefore, the grouping number range for first tape is 1 to 30 and the grouping range for the second tape is 1 to 13.

The RE tar files contain a group of RE's such that one RE tar file contains at most ten sub directories. Each sub directory in the tar file contains all the related data used as input to create at least one of the SIA data file. Each sub directory name follows the convention:

ENTRY/

where:

ENTRY = the SDB entry number

Each RE file has been named following these guidelines:

Positions 1-4 Platform:

n11_ = NOAA N_11
n12_ = NOAA N_12
f10_ = DMSP F_10
f11_ = DMSP F_11
g04_ = GMS-4
m04_ = METEOSAT-4

Positions 5-8 Type of file:

001_ = satellite data channel 1
002_ = satellite data channel 2
...
...
005_ = satellite data channel 5
lat_ = latlon data
ang_ = angles data
mcf_ = cloud mask data
sdb_ = SDB information file

Positions 9-12 Area of data:

eas_ = East Asia Area (EASA)
can_ = Canada Area (CANA)
cns_ = Central and Northern South America Area (CNSA)
emd_ = Easter Mediterranean, Desert Area (EMDA)

Positions 13-16 Julian day:

081_ = Julian day 081 etc. ...

Positions 17-18 Hour:

00 = hour of the data

Positions 19-22 Extension:

.dat = raw data
.tif = tif formatted data

Examples:

f10_001_eas_150_14.tif
f10_002_eas_150_14.tif
f10_lat_eas_150_14.tif
f10_ang_eas_150_14.tif
f10_mcf_eas_150_14.tif
f10_sdb_eas_150_14.tif

****

How do I generate a listing of the contents of the release tapes ?

To generate a listing of the contents of the release tapes utilize the following UNIX "list_tar" script:

```
>#! /bin/csh -f -x
>
>while ($status == 0)
>
>       echo " "
>       tar vtf /dev/sometapedrive
>       if (($status != 0) && ($status != 3)) then
```

```
>               exit
>       endif
>end
>exit
```

NOTE: replace "sometapedrive" with the proper tape drive identifier.

This script should be created with a basic UNIX text editor (i.e. vi,emacs) and given the filename of list_tar. Once the script has been created the following command can be used to direct the output of the "list_tar" script to a file:

```
>       %list_tar > tape1.contents
```

NOTE: Generating the tape contents listing will help supplement this document.

****

What are related data items ?
What is the SDB entry number ?
What are related entries ?

The SDB registration process is a process that automatically places descriptive data items about a satellite scan into the SDB. The SDB registration process allocates a group of unique entry numbers to be used as place holders for all of the related data items for a given satellite scan. The related data items consists of satellite, latitude/longitude, angles (Geometry) and product(cloud mask) data. As an example, if a DMSP F_11 scan was to be registered in the SDB, the registration process would request for a group of five contiguous entry numbers(i.e. 1001-1005). These five entry numbers would be used as place holder for the following related data items:

1001	f11 visible channel
1002	f11 infrared channel
1003	latitude/longitude data
1004	angles(geometry) data
1005	product data

The "SDB entry number" is the first entry number of the group of entry numbers provided by the registration process. The first entry number is used to "key" into the related data items for that group. In the example provided above the SDB entry number would be 1001.

The release process uses the SDB entry number in each group to logically divide the data into separate directories (i.e. the directory name is first SDB entry number for each group of entry numbers). Using the example provided above the directory named "1001/" contains all the related data items for that group (i.e. the directory contains the data for entry 1001 through entry 1005).

To build a SIA it is necessary to use as input, related data items from one or more satellite scans and/or satellite platforms. The SDB entry number is used to keep track of all inputs to the SIA. The list of related entries are given as SDB entry numbers.

****
How do I get a particular SIA data set ?

You must use the UNIX tar utility to extract the data from the tape. By using the tar utility you may extract individual files or the entire directory. Use a tape contents list generated using the "list_tar" script described above to determine where to position the tape and then use the appropriate tar command to extract the files you want. For example, if you want to extract all of the SIA data files from the first and second tar files, the following commands could be used:

```
% pwd
/users/smith
% mkdir data
% cd data
% tar xvf /dev/rst8 994071
% tar xvf /dev/rst8 994072
```

Upon completion all of the SIA data for day 071 would reside in directory /users/smith/data/994071 and all the SIA data for day 072 would reside in directory /users/smith/data/994072.

****
What is the SDB information file ?

The SDB information file is a text file containing selected SDB record items that help describe the actual data. The SIA SDB information file shows what data went into creating the SIA by listing the related entries. The RE SDB information file lists information about the satellite images, the latlon file, the angles file and the product file(s).

The following is an example SIA SDB information file:

```
[IA]
ZULU_YYJJJ:=94071                        : Year, Julian day of SIA
ZULU_HH:=10                              : Hour of SIA
ROI:=EMD                                 : Region of Interest
NUM_RELATED_LAYER:=3                     : Number for related entries
RELATED_LAYER_1:= 4148                   : 1st related SDB entry number
RELATED_LAYER_2:= 7199                   : 2nd related SDB entry number
RELATED_LAYER_3:= 8988                   : 3d related SDB entry number
TDISK:=SDB_Int:
TDIR:=[SERCAA.DATA.994071]
FILE_IA_1:=ALL_IAN_EMD_071_08.Dat        : SIA file name
SDB_SET:=MAR94                           : Set identifier March of 1994
```

The following is an example RE SDB information file:

```
[SATIMG]
SAT_CODE:=16                                                 : Satellite code
ZULU_YYJJJ:=94071                                            : Year, Julian day of scan
ZULU_HHMMSS:=82252                                           : Time of scan
NUM_LINES:=1375                                              : Number of lines
ELEM_PER_LINE:=409                                           : Elements per line
BYTES_PER_ELEM:=1                                            : Bytes per element
7199:=AVH$005:[SERCAA.DATA.994071]N11_001_EMD_071_08.TIF     : Channel 1 file
7200:=AVH$005:[SERCAA.DATA.994071]N11_002_EMD_071_08.TIF     : Channel 2 file
7201:=AVH$005:[SERCAA.DATA.994071]N11_003_EMD_071_08.TIF     : Channel 3 file
7202:=AVH$005:[SERCAA.DATA.994071]N11_004_EMD_071_08.TIF     : Channel 4 file
7203:=AVH$005:[SERCAA.DATA.994071]N11_005_EMD_071_08.TIF     : Channel 5 file
```

[LATLON]	
LL_REC_LEN:=204	: Record length in bytes
LL_LINE_INTERVAL:=1	: Sub-sample line interval
LL_ELEM_INTERVAL:=8	: Sub-sample element interval
LL_ELEM_PER_LINE:=51	: Latlon pairs per line
LL_FILE:=AVH$005:[SERCAA.DATA.994071]N11_LAT_EMD_071_08.DAT	: latitude/longitude file
[ANGLES]	
ANG_REC_LEN:=612	: Record length in bytes
ANG_LINE_INTERVAL:=1	: Sub-sample line interval
ANG_ELEM_INTERVAL:=8	: Sub-sample element interval
ANG_ELEM_PER_LINE:=51	: Angles triplets per line
ANG_FILE:=AVH$005:[SERCAA.DATA.994071]N11_ANG_EMD_071_08.DAT	: Angles file
[PRODUCT]	
7206001:=sdb$prd:[SERCAA.DATA.994071]N11_MCF_EMD_071_08.TIF	: Cloud mask file

****

How do I know which RE data went into a particular SIA ?

There are two ways to determine which RE data sets went into a particular SIA. The first way is reference the SIA SDB information file. Each "RELATED_LAYER" listed is a reference, by SDB entry number, to the RE data. Use the referred SDB entry number to retrieve the related data from the RE data tape.

For example, refer to the above SIA SDB information file. The "RELATED_LAYERED_1:=4148" line implies that SDB entry number 4148 and the related data items for entry 4148 (along with SDB entry numbers 7199 and 8988) were used to create "ALL_IAN_EMD_071_08.Dat".

The second way is to read the header information from the SIA file (Please refer to the DATA_DESCRIPTION).

****

How do I get the RE data files ?

Once you have examined the SIA SDB information file and you have identified the related entry numbers, you must use the UNIX tar utility to extract the data from the tape. By using the tar utility you may extract individual files or the entire directory. Use a tape contents list generated using the "list_tar" script to determine where to position the tape and then use the appropriate tar command to extract the files you want. For example, if you want to extract the RE data files for entry 7199 from the group 1 tar file, the following commands could be used:

```
        % pwd
        /users/smith
        % mkdir data
        % cd data
        % tar xvf /dev/rst8 1/7199
                            ^  ^
                            ||  ||
 grouping number-------     ------entry number
```

Upon completion of this command all of the RE data related to SDB entry number

7199 would reside in directory /users/smith/data/7199.

****

For the following question please refer to the example SDB information files
as needed.

****

What is the format of the satellite data and how do I access it?

The dimensions of the satellite data are defined by the three parameters, NUM_OF_LINES, ELEM_PER_LINE and BYTES_PER_ELEM . To access the data use the following logic.

If the file extension is ".dat"
then use the appropriate C or FORTRAN read statements.

If the file extension is ".tif"
then use a TIF reader or TIF library (you may view the
images by using the public domain application, XV).

For a detailed explanation, refer to Appendix B.

****

What is the format of the latlon data and how do I access it?

The latlon data are sub-sampled. The dimensions are defined LL_LINE_INTERVAL, LL_ELEM_INTERVAL and LL_ELEM_PER_LINE. LL_ELEM_PER_LINE defines the number of longitude/latitude pairs per line. Each pair is four bytes (two bytes lon, two bytes lat). To access the data use the appropriate C or FORTRAN read statements.

For a detailed explanation, refer to Appendix B.

****

What is the format of the angles data and how do I access it?

The angles data are sub-sampled. The dimensions are defined by ANG_LINE_INTERVAL, ANG_ELEM_INTERVAL and ANG_ELEM_PER_LINE. ANG_ELEM_PER_LINE
defines the number of triplets (satellite-zenith/solar-zenith/azimuth) per line. Each item in the triplet is a float data type. To access the data use the appropriate C or FORTRAN read statements.

For a detailed explanation, refer to Appendix B.

APPENDIX B

AIMS SATELLITE DATABASE PRESENTATION

Current AIMS Database

- Design based on requirements from SERCAA Phase I research and development program

  - Program objective was to provide a multispectral multisource global cloud analysis capability for use in determining the radiative and hydrological effects clouds have on climate and global change

- A satellite image database

SERCAA Data Processing Requirements

- Process sensor data from multiple satellite platforms including DMSP/OLS, NOAA/AVHRR, GOES/VAS, METEOSAT/VISSR, and GMS/VISSR
- Series of processes to analyze and integrate data into a single cloud analysis product
  - Ingest
  - Total cloud
  - Cloud typing, layering, and heights
  - Cloud integration
- Collect data from three geographically defined test-bed areas collectively covering just over half the northern hemisphere
  - Last 10 days of every other month
  - March 1993 through September 1995

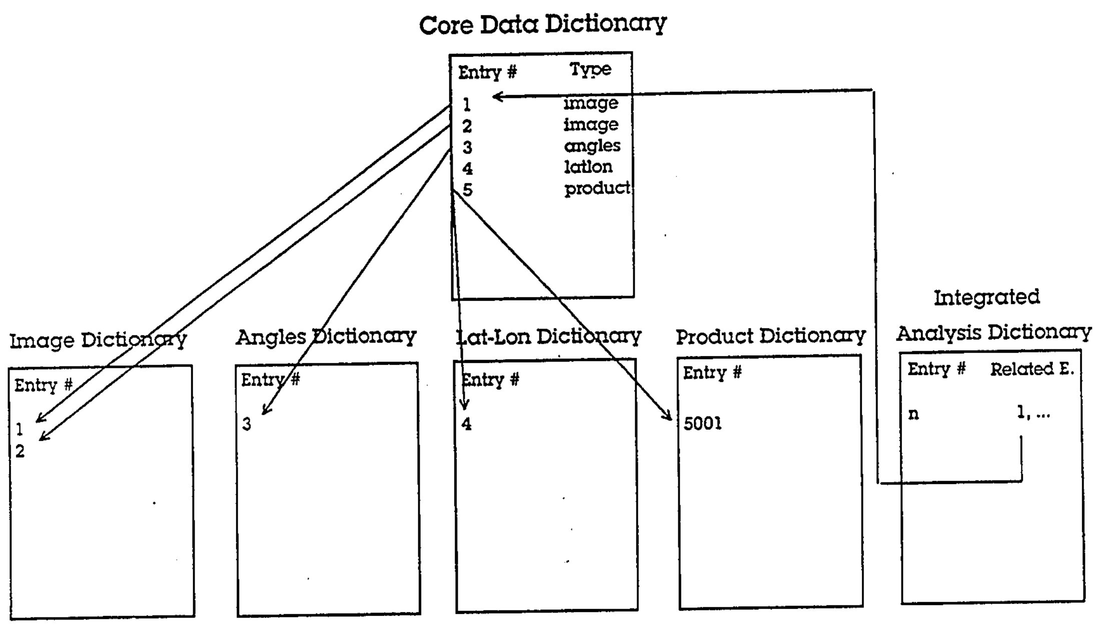
SERCAA Database Functional Requirements

- Ingest satellite data (scan line format) from multiple platforms into the database
- Manage data from multiple 10-day collection periods
- Ability to create, read, modify, and delete database entries
- Interactive and programmed access to the database
- Programmed access to data files

Database Characteristics

- Manages metadata contained in a set of interrelated data dictionaries
- Data dictionaries and metadata structures inherited from TACNEPH program
  - Organization based on the relational model
  - Data dictionaries implemented as index files on VMS providing keyed access to records
  - Key used for TACNEPH/SERCAA is the entry number
- Programmed access to the data dictionaries using subroutines to create, read, modify, and delete entries
- Interactive access using general-purpose programs (SDB, Examine, and Update)
- Each 10-day data save is managed using a unique set of data dictionaries (user selectable)

Representation of the Relational-Hierarchical Set of Data Dictionaries That Comprise the SERCAA Database



Improvements to the TACNEPH Database

- Elimination of fragmented data dictionary files improved I/O performance
- File open procedures were changed to allow multiple concurrent access to the data dictionaries by both readers and writers
- A more efficient procedure was developed to manage entry numbers

Extensions to the TACNEPH Database

- New Core Data Dictionary Query Routine

  - Addressed the need to access database records based on date-time, satellite platform, and location

  - Required the creation of a new secondary core data dictionary sorted on the above fields

  - The result of a query would produce entry number(s) that the user would use to access entries in the primary core data dictionary

  - Incorporated into the general-purpose interactive query program (SDB)

Sample Output from the SDB Program Utilizing the Core Data Dictionary Query Routine

```
$ SDB_Select JUL93
$ SDB/Query=Image/Time=93209/Sat=NOAA_12
```

Entry	Type	Date	Time	Area	Sat	Channel	Bkup	Pixels	Lines	Resolution
1705	IMG	93208	235005	EAS	NOAA_12	AVHRR-1	Yes	409	2098	4.00
1706	IMG	93208	235005	EAS	NOAA_12	AVHRR-2	Yes	409	2098	4.00
1707	IMG	93208	235005	EAS	NOAA_12	AVHRR-3	Yes	409	2098	4.00
1708	IMG	93208	235005	EAS	NOAA_12	AVHRR-4	Yes	409	2098	4.00
1709	IMG	93208	235005	EAS	NOAA_12	AVHRR-5	Yes	409	2098	4.00
1721	IMG	93209	092842	EAS	NOAA_12	AVHRR-1	Yes	409	1652	4.00
1722	IMG	93209	092842	EAS	NOAA_12	AVHRR-2	Yes	409	1652	4.00
1723	IMG	93209	092842	EAS	NOAA_12	AVHRR-3	Yes	409	1652	4.00
1724	IMG	93209	092842	EAS	NOAA_12	AVHRR-4	Yes	409	1652	4.00
1725	IMG	93209	092842	EAS	NOAA_12	AVHRR-5	Yes	409	1652	4.00

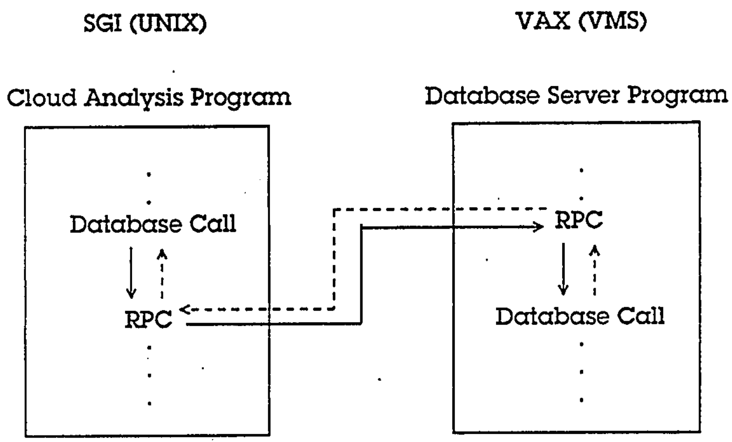
Extensions to the TACNEPH Database (Cont'd)

- New Satellite Data Access Routine
  - Addressed the need to abstract the details of file I/O to a level below the application program so that the programmer could concentrate on data processing
  - Access to an entire image at one time
  - Access to multiple scan lines of an image
  - Access to geographic subsets within an image scene
  - Support multiple file formats (TIFF, RAW)

Extensions to the TACNEPH Database (Cont'd)

- Database access from UNIX platforms
  - Addressed the need for SGI workstations performing cloud typing, layering, and integration processing to access the database and satellite data
  - Remote Procedure Call (RPC) used in a client-server configuration

Use of RPC by SERCAA for Processing Levels 3 & 4



Benefits:

- No need to port database code
- Network communications hidden from the programmer
- Bulk of RPC code generated with the RPCgen tool
- Byte swapping and floating point conversion is automatic

Needed Improvements to the Current Database Capability

- More robust client-server solution for distributed access to the database
- A flexible, GUI-based general-purpose application for making database queries, modifications, and deletions
- Extensible
- Manage satellite data in addition to the metadata?
- Data quality control
- Minimize/eliminate maintenance of include files
- Minimize/eliminate kludging of database resources
- Ability to dynamically modify record structures
- Ability to pose queries across dictionary boundaries

Database Management Systems Benefits

AIMS users spend large amounts of time with issues concerning data **handling** and manipulation

This time would be much better spent on data **analysis** issues

Data handling issues include such details as:

- Searching for suitable data sets of interest
- Locating data and making available to application programs
- Geo-location, co-location, and calibration of data sets and ensuring data quality and data integrity
- Numerous network and hardware compatibility issues
- Visualization

Slide No. 1

DBMS Key to Solving Problems of Info Management

DBMS provides a level of abstraction between physical data storage and logical data structure

Data managed by DBMS is device independent

Users and their application programs need not be concerned with physical storage structure and storage device

DBMS software supports the of network transparency

Presents consistent interface for the sharing of data and allows others to access the database directly

Slide No. 2

Data Worthy of Putting in Database

Virtually any data set can be managed by the DBMS including

Remotely sensed data

Balloon and Aircraft generated data

Ground observations

...and just about any thing else

What are the major attributes upon which we would like to logiacally relate data sets? (i.e. time and location)

Slide No. 3

With a Time and Location Originization of Data Sets

For example

GET ALL DATA
WITHIN regionX
BETWEEN time1 AND time2
FROM PLATFORM goes8, noaa12, aircraft1...

Such a query not possible now

Slide No. 4

Database Architecture

There are several ways one might manage data using a DBMS

One of the central choices is how much of the data should be inserted into the database

Do we store meta data only and pointers to the larger data sets Otherwise, database could become slow as it grows in size

Extra work to ensure data integrity since this must be programmed at the DBMS system level and is no longer automated

Slide No. 5

Database Architecture

OR

Do we store full size data sets in the database (BLOBS or Q-TREEs) taking advantage of all built-in DBMS integrity constraints

This allows for complex queries and data retrieval operations with little or no DBMS system programming

Could be slow but do to nature of database (no OLTP) should be feasible

Slide No. 6

Data Format

Standardized data file formats allow for the physical grouping of logically related data and enhance portability

Data file formats include:

HDF, CDF, NITF, SDTS, GeoTIFF, etc...

Each format has benefits and deficiencies

Numerous software libraries available to support these formats

Can support several but evaluation criteria include:

Efficiency, Generality, How widely accepted, Extensable

What format(s) make the most sense for us?

Slide No. 7

DBMS Access Methods

Access Method	Availability	Required Experience
Oracle Data Browser	immediate	GUI interface, little or none
WWW Page Interface	to be developed	GUI interface, little or none
SQL Command Line	immediate	Moderate SQL experience
Embedded SQL or API	to be developed	Used by app. programmers

Slide No. 8

Oracle Data Browser

Point and Click interface

Used in house for

- Formulating complex queries to check for existence of data
- Reduces need to be SQL literate
- Included with DBMS software, i.e. no development required
- Immediate availability

Slide No. 9

WWW Page Interface

Point and Click interface

Used in house and **elsewhere** for

- viewing data sets as images
- formulating complex data queries
- downloading image sets or some other product

Would require development and maintenance but would leverage much existing technology

Allow world wide community to see (buy?) products produced by PL/AER (Important to be recognized and acknowledged)

Slide No. 10

SQL Command Line Interace

Command line driven utility program (not point and click!)

Used in house for

- Formulating complex data queries
- Data table creation and maitanence
- Primarily used by DBA or technical user with need
- More powerful and flexible than GUI interface
- Scripting language built-in suitable for performing automated or complex tasks
- Requires familiarity with SQL and many database concepts

Slide No. 11

Embedded SQL and API

SQL commands inserted directly into application program or inserted inderectly through application program interface

Used by application programmers for:

- Retrieval of data from database directly into application
- Embedded SQL product part of database - requires database knowledge
- API calls look like standard library call - requires no database knowledge but libraries would need to be developed

Slide No. 12